Dynamic Persuasion and Strategic Search

Yunfei (Jesse) Yao[*]

(University of California, Berkeley)

June 24, 2022

Latest Version: *Click Here*

Abstract

Consumers frequently search for information before making decisions. Since their search and purchase decisions depend on the information environment, firms have a strong incentive to influence it. This paper endogenizes the consumer's information environment from the firm's perspective. We consider a dynamic model where a firm sequentially persuades a consumer to purchase the product. The consumer only wishes to buy if the product is a good match. The firm designs the information structure. Given the endogenous information environment, the consumer trades off the benefit and cost of information acquisition and decides whether to search for more information. Given the information acquisition strategy of the consumer, the firm trades off the benefit and cost of information provision and determines how much information to provide. This paper characterizes the optimal information structure under a general signal space. We find that the firm spreads information provision over multiple periods if and only if the prior belief for a good match is high. In equilibrium, if the search cost for the consumer is high, the firm will design the information such that the consumer will be certain that the product is a good match and will purchase it right after observing a positive signal. If the search cost is low, the firm provides noisy information such that the consumer will be uncertain about the product fit but still makes the purchase right after observing a positive signal. We also compare the profit-maximizing and the efficient (maximizing social welfare) information structures. The profit-maximizing strategy is efficient if and only if the search cost is high.

1 Introduction

   With the rapid proliferation of digital technologies and information channels, it is increasingly common for consumers to seek detailed information before making a decision. More information can lead to less uncertainty and improve decision-making. Since consumers' search and purchase decisions depend on the information environment, firms have strong incentives to influence it. Thus, the consumer faces endogenous information environment. For example, a consumer considering purchasing a pair of shoes may search on the internet to find out whether or not the item matches his needs.[1] The seller can influence consumers' search and purchase decisions through various methods. She can bid for search advertising spots to persuade the consumer. The content can be informative, showing features about the item, or persuasive, without detailed information. By spending more money for advertising, the seller can communicate information to potential buyers more frequently. Even if the consumer does not make a purchase right after seeing an ad, the seller can keep persuading the consumer through retargeting. At the same time, the consumer spends time and effort to search. He will only search for more information if he anticipates enough gain from it.

   Such communications are also ubiquitous offline. For instance, an individual looking for a new car often visits a local dealership to talk with the dealer. The dealer will highlight various characteristics of the car. If the consumer is interested, the dealer often offers him a test drive to have a better sense of the car. This is time-consuming, in contrast to the ease of searching for information about a product on the internet, but they are often not sure whether they will like a product based solely on the information online. In the case of car-buying, test driving and then haggling with the dealer can take several hours, but people who purchase the car after visiting the dealership usually are sure that they like the car. Motivated by these examples, this paper endogenizes the consumer's information environment from the firm's perspective.

   To better understand the above scenarios, we consider a dynamic model where a receiver makes a binary decision between action G and B. There are two states, good (g) and bad (b). A sender always prefers G and sequentially persuades the receiver to take that action. In contrast, the receiver only wishes to take action G if the state is good. Neither the sender nor the receiver knows the state initially but have a common prior belief about it. The receiver can incur costs to search for more information about the state. The updated belief helps him make decisions. If the receiver observes a negative signal, he knows that the state is less likely to be good and will not take action G without the arrival of new information.

---

[1] We refer to the information provider (seller) as "she" and the decision-maker (consumer) as "he" throughout the paper.

If the receiver observes a positive signal, he knows that the state is more likely to be good, and the expected payoff of taking action G increases.

The sender designs the information structure. Given the endogenous information environment, the receiver trades off the benefit and cost of information acquisition and decides whether to search for more information. The receiver is forward-looking and forms rational expectations of the sender's strategy. The sender could incur higher costs to provide more information so that the receiver would be more likely to take action G. Given the information acquisition strategy of the receiver, the sender trades off the benefit and cost of information provision and determines how much information to provide. Therefore, the receiver and the sender simultaneously trade off the benefit and cost of information acquisition/provision.

This paper characterizes the optimal information provision strategy of the sender and the optimal information acquisition strategy of the receiver. There are two periods in the main model. In each period, the sender chooses the information structure, and the receiver chooses whether to search for more information or make a decision. We later extend the two-period model to an infinite-period model and show that the main insights extend to the richer model. Instead of looking at specific parameters of the search environment, we study the design of the information environment under general signal space and characterize the optimal information structure among all feasible information policies. We develop an optimal control method to solve the sender's information design problem, because the widely-used concavification method due to Kamenica and Gentzkow (2011) cannot solve it when the receiver's particiaption is strategic.

In equilibrium, the sender uses one-shot experiments, which induce the receiver to take action G immediately upon observing a positive signal. This way, the sender saves the expected search time and does not need to compensate the receiver for a higher expected search cost. The sender spreads information over multiple periods if and only if the prior belief about the good state is high. The sender extracts all the surplus from the receiver when she provides information in both periods, while she may leave some surpluses to the receiver when she provides information in only one period. The sender has an incentive to spreads information provision over multiple periods. However, the longer expected search time of the receiver would discourage him from searching if the likelihood of getting a strictly positive ex-post payoff is low. Hence, the sender only smooths information if the prior belief is high. If the search cost for the receiver is high, the sender will design the information such that the receiver will be certain that the state is good and will take action G right after observing a positive signal. If the search cost is low, the sender provides noisy information such that the receiver will be uncertain about the state but still takes action G right after observing a positive signal.

We compare the profit-maximizing structure with the efficient (social welfare maximizing) information structure. When the search cost is high, the optimal strategy of the sender also maximizes social welfare. When the search cost is lower, the two information structures are different because the sender does not internalize the receiver's welfare.

In the main model, the sender can commit to the current-period information structure but cannot commit to the information structure in the next period. In the extension, we study the implications of dynamic commitment power. We find that the ability to commit to future strategies strictly benefits the sender when the search cost is high, while the benefit vanishes as the search cost approaches zero.

We also consider the implications of discounting. When the search cost is high, it is difficult to convince the receiver to participate. Providing enough information to persuade the receiver to search dominates the force of discounting. So, the sender's optimal strategy does not depend on the discount factor. When the search cost is low and the players become less patient, the present value of the second-period profit decreases and the first-period participation constraint becomes tighter. The sender has a stronger incentive to convert the receiver early. So, she provides less information in the second period and more information in the first period. Moreover, we consider asymmetric discount factors in a model where the sender is patient while the myopic receiver only cares about the current-period payoff. We find that both players are worse off if the receiver is myopic rather than forward-looking. This implies that the sender should try to better inform the receiver of the possibility of multi-period information revelation and gradual learning.

Lastly, we extend the two-period model to an infinite-period model and show that the main insights extend to the richer model. When the search cost approaches zero, the ability to smooth the information over more extended periods is valuable for the sender. In that case, she could obtain the equilibrium payoff as if the persuasion cost were zero. We also find that the receiver will decide within a bounded number of periods.

1.1 Related Literature

There is a large stream of literature on optimal information acquisition. In particular, consumer search has raised growing interest both theoretically (Stigler 1961, Weitzman 1979, Moscarini and Smith 2001, Branco et al. 2012, Ke et al. 2016, and Ke and Villas-Boas 2019) and empirically (Hong and Shum 2006, Kim et al. 2010, 2017, Seiler 2013, Honka 2014, Ma 2016, Chen and Yao 2017, Honka and Chintagunta 2017, Seiler and Pinna 2017, Ursu et al. 2020, Moraga-González and Wildenbeest 2021). In the above papers, the information environment is exogenous. Because the strategy and the outcome depend on the information

structure, other payoff-relevant parties (e.g., the seller) have strong incentives to influence it. We take this into account by having the firm endogenously determine the information environment of the consumer.

Several papers study consumers' endogenous information acquisition. The consumer chooses both the search rule and the information environment. In Zhong (2019), the decision-maker gradually gathers information about one product. Poisson learning is optimal for him. In Guo (2021), the consumer sequentially search for information about multiple products and determines how much to evaluate each product. In our paper, by contrast, the seller endogenizes the search environment.

A few papers also investigate the design of the search environment from the firm's perspective. In Dukes and Liu (2016), consumers choose the breadth and depth of the search. The platform designs the search environment by selecting the search cost. Villas-Boas and Yao (2021) consider the optimal retargeting strategy of the firm which advertises to consumers who have a high likelihood of considering the firm's product. By advertising, the firm increases the frequency of consumers' learning information and the ability to track consumers. In Zhong (2022), the platform recommends relevant sellers based on match values and prices to consumers. The platform designs the search algorithm by picking the match precision and the relative importance between prices and match values. Comparing the welfare outcomes among information structures emphasizing different vehicle characteristics under the counterfactuals of their structural model, Gardete and Hunter (2020) find that emphasizing the vehicle's history and obfuscating price information improves both consumer and firm welfare. In related literature on choice overload, the seller determines the amount of information provided to the consumer (Kuksov and Villas-Boas 2010, Branco et al. 2016). The decision of the seller affects the consumer's search cost.

While the above papers are related to our paper, there are substantial differences. Instead of looking at specific parameters of the search environment, we study the design of the information environment under a general signal space. We characterize the optimal information structure among all feasible information policies. In the existing literature, the consumer either fully observes the value of a product/attribute or gets a noisy signal from an exogenously specified distribution (e.g., a normal distribution). In contrast, in our paper, the firm determines the distribution of the signal. The optimal information structure can be asymmetric and may not correspond to standard distributions.

We use a belief-based method of modeling information provision, first introduced by Aumann and Maschler (1995) and Kamenica and Gentzkow (2011) in the Bayesian persuasion and information design literature.[2] The sender picks a mean-preserving spread of the

---

[2] See Bergemann and Morris (2019) for a survey.

prior as the posterior belief, which simplifies the analysis. Some papers have studied the persuasion problem when either the receiver or the sender incurs costs. For example, Wei (2021) considers a static persuasion problem in which a rationally inattentive receiver incurs information processing costs. Gentzkow and Kamenica (2014) extend the widely-used concavification approach of Kamenica and Gentzkow (2011) to the setting where the sender's cost is posterior-separable. However, the concavification approach cannot be used if the receiver's particiaption is strategic or the sender's cost is not posterior-separable. So, we instead develop an optimal control method to solve the sender's information design problem. The constraints of the non-linear program of the sender consist of multiple variables, making the optimization problem challenging. To make the problem tractable, we transform the program into a set of constrained programs, each with one constrained variable. We then select the global solution by comparing the local solutions of each program.

While some recent papers extend the static setting of Bayesian persuasion to the dynamic one, it is usually costless for the sender to provide information (Ely 2017, Ely and Szydlowski 2020, Iyer and Zhong 2021, Orlov et al. 2020). One exception is Che et al. (2020), in which the sender performs costly experimentation to persuade the receiver to take a particular action. The flow cost of providing the information is constant, and the type of experimentation available to the sender is limited. In our paper, the sender can speed up the receiver's learning by incurring a higher cost. So, we can investigate the optimal tradeoff of the benefit and cost of information provision by the sender.

The remainder of the paper is organized as follows. Section 2 presents the main model. Section 3 characterizes the optimal information provision strategy and the equilibrium outcomes. Section 4 characterizes the efficient information provision strategy and summarizes the information distortion when the sender rather than the social planner designs the information structure. Section 5 includes several model extensions. Section 6 concludes.

2　The Model

2.1　States, Actions, and Payoffs

There are two players, a sender and a receiver, and two states, good (g) and bad (b). The receiver ultimately makes a binary decision between G and B. The sender wishes to persuade the receiver to take action G regardless of the state, while the receiver wishes to match the decision with the state (taking action $G(B)$ when the state is $g(b)$). The payoffs of the decision for the players are the following:

(sender payoff, receiver payoff)	action G	action B
state g	(p, v_g)	$(0,0)$
state b	(p, v_b)	$(0,0)$

The sender earns a positive payoff, $p > 0$, if the receiver takes action G. The receiver's payoff is positive if he takes action G when the state is g, $v_g > 0$, and negative if he takes action G when the state is b, $v_b < 0$. Both players get zero payoff if the receiver takes the action B (which could be thought of as an outside option). We assume without loss of generality that $v_g = 1 + v_b$.[3] Neither the sender nor the receiver knows the state initially but have a common prior belief about it, $\mu_0 := \mathbb{P}(\text{the state is } g) \in (0, 1)$. In each period $t \in \{0, 1\}$, the sender determines and commits to the information structure of the experiment but cannot commit to the experiment in the future. The receiver can search for information (action S) before deciding. The information acquisition is costly but helps the receiver make better decisions. If the receiver chooses to search, he observes an experiment that reveals some information about the state. The sender does not observe the realization of the signal but can infer it from the receiver's action.[4] The players update the belief after the realization of the experiment. The game ends whenever the receiver makes a decision (G or B). Figure 1 illustrates the timing of the game.
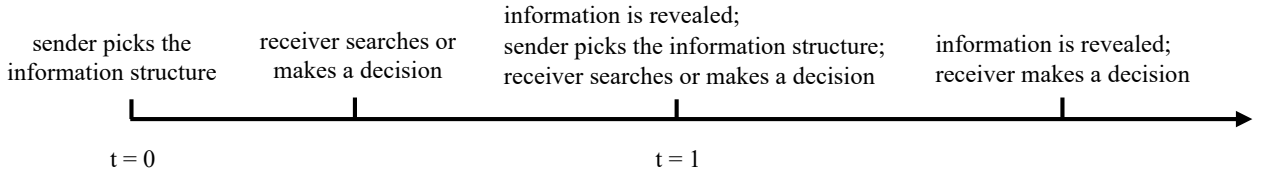

Figure 1: Timing of the Game

We model the experiment by a binary signal $s \in \{0, 1\}$.[5] $\mathbb{P}[s = 1|g]$ and $\mathbb{P}[s = 1|b]$ uniquely determines the signal. We order the value of the signal such that $\mathbb{P}[s = 1|g] > \mathbb{P}[s = 1|b]$. Hence, $s = 1$ corresponds to a positive signal and $s = 0$ corresponds to a negative signal. Analogous to Proposition 1 of Kamenica and Gentzkow (2011), we can work with mean-preserving posterior beliefs rather than the specific signal structure to simplify the

---

[3] To see that assuming $v_g = 1 + v_b$ is WLOG, consider the following normalization. Let $v'_g = \frac{v_g}{v_g - v_b}, v'_b = \frac{v_b}{v_g - v_b}, p' = \frac{p}{v_g - v_b}$. Then, $v'_g - v'_b = 1$. We make this assumption to simplify the analysis and the presentation.

[4] From Proposition 1, one can see that the sender can perfectly infer the realization of the signal from the receiver's action under the optimal signal structure. So, the receiver does not have private information.

[5] Since there are only two states, more than two signal values will not benefit the sender.

analysis. Specifically, the existence of a binary signal s such that $\mathbb{P}[s = 1|g] > \mathbb{P}[s = 1|b]$ is equivalent to the existence of a binary-valued posterior belief whose expectation is equal to the prior.[6]

Denote the belief at the beginning of each period by μ_t. At each period, with probability λ_t, the receiver observes a positive signal and the belief increases to $\bar{\mu}_t$. We refer to λ_t as the probability of a positive signal and $\bar{\mu}_t$ as the belief after observing a positive signal. With probability $1 - \lambda_t$, the receiver observes a negative signal, and the belief decreases to $\underline{\mu}_t$. We refer to $\underline{\mu}_t$ as the belief after observing a negative signal.

We assume there is no discounting in the main model, and discuss the implication of discounting in the extension. The receiver incurs a flow cost of c per period of search. The sender's cost of information provision depends on the probability of a positive signal, $K = K(\lambda)$. The sender incurs a higher cost to provide information structures with more frequent positive signals. Throughout this paper, we refer to the likelihood of a positive signal as the amount of information. More information means more frequent positive signals. We make the following assumptions about the cost incurred by the sender.

Assumption 1. $K(\cdot) \in \mathcal{C}^2(\mathbb{R}_+), K'(\lambda) > 0, K''(\lambda) > 0, \ K(0) = 0, \ \lim_{\lambda \to 1^-} K'(\lambda) = +\infty, \ \lim_{\lambda \to 0^+} K'(\lambda) = 0.$

This assumption means that it is relatively cheap for the sender to provide a little information. The marginal cost of the sender increases at an increasing rate when the amount of information provided increases.

The total payoff for each player is the payoff of the decision net of the information provision/acquisition costs. The receiver is forward-looking and forms rational expectations about the sender's strategy in the future. We assume that $\mu_0 v_g + (1 - \mu_0) v_b < 0 \Leftrightarrow \mu_0 < -v_b$. So, the receiver would never take action G without searching. Otherwise, the sender would provide no information, and the receiver always would take the desired action. We also assume that the search cost is not too high, $c < v_g$. Otherwise, the receiver would never search.

2.2 Applications

Our model can be applied to many settings where a sender wishes to persuade a receiver to take a particular action while the receiver wishes to match the action with the state and can strategically gather more information before making the decision. We illustrate two applications of the model in this section.

---

[6] In the appendix, we state this result formally and provide proof.

2.2.1 Product Sales

A seller offers a product with unit demand. The price is p, and the marginal cost is m. The product may be a good match or a bad match with the buyer. Neither the buyer nor the seller knows the state initially but they have a common prior belief. The product is of zero value to the buyer if the state is bad and is of value v to the buyer if the state is good. The buyer decides whether or not to purchase the product. The payoffs of the decision for the players are the following:

(sender payoff, receiver payoff)	purchasing	not purchasing
state good match	$(p-m, v-p)$	$(0,0)$
state bad match	$(p-m, -p)$	$(0,0)$

The buyer could incur costs to search for more information about the product. The updated belief helps him make better purchasing decisions. If the buyer observes a negative signal, he knows that the product is less likely to be a good match and avoids wasting money on it without the arrival of new information. If the buyer observes positive news, he knows that the product is more likely to be a good match and the expected valuation increases. The buyer gets a positive surplus upon purchasing the good if the expected value is higher than the price.

2.2.2 New Drug Launches

A pharmaceutical company develops a new drug and tries to get approval from the regulator (e.g., the FDA) by designing and conducting tests to convince the regulator that the drug is safe. If the regulator approves it, the firm gains an expected profit of p. The regulator gains a positive utility if the drug is safe and a negative utility if it is unsafe. The payoffs of the decision for the players are the following:

(sender payoff, receiver payoff)	approval	disapproval
state safe	$(p, v_g > 0)$	$(0,0)$
state unsafe	$(p, v_b < 0)$	$(0,0)$

Monitoring and examining the test is costly for the regulator. So, the regulator would only investigate the test if both the likelihood that the result is positive and the precision of

the test are high enough. Even if a single test fails to convince the regulator, the firm could offer another test, hoping that the result would finally convince the regulator to approve the drug.

2.3   Strategies and Equilibrium Concepts

Since the belief is common knowledge and there is no private information, we consider the sender-preferred subgame perfect equilibrium, as in Kamenica and Gentzkow (2011). If multiple actions $(B, G,$ and $S)$ give the receiver the same expected payoff, we assume that the receiver chooses an action that maximizes the sender's expected payoff. We also assume that the sender prefers to give the receiver more surplus in the first period if more than one equilibrium lead to the same expected sender's payoff. A perturbation of very little discounting justifies this assumption.

2.3.1   Equilibrium of the Subgame

In the second period, the sender and the receiver play the subgame. Since it is costly for the sender to provide information, the sender would either give no information or provide enough information such that the receiver searches and will take action G if a positive signal arrives. We illustrate the belief evolution of the subgame in Figure 2. Also, the distribution of the belief induced by the experiment should be a mean-preserving spread of the initial belief: $\mathbb{E}[\Delta\mu] = 0$. In sum, the sender either does not provide information and obtains zero payoffs or takes into account the following constraints when designing the information structure:

(1) participation constraint:

$$\lambda_1[\bar{\mu}_1 v_g + (1 - \bar{\mu}_1)v_b] = \lambda_1(\bar{\mu}_1 + v_b) \geq c \tag{IR_1}$$

(2) feasibility constraint:

$$\lambda_1\bar{\mu}_1 + (1 - \lambda_1)\underline{\mu}_1 = \mu_1 \tag{F_1}$$

If the sender provides information, the constrained program of the sender is:

$$\max_{\lambda_1, \bar{\mu}_1} -K(\lambda_1) + p\lambda_1 \tag{P_1}$$

$$\text{s.t. } (IR_1), (F_1), \lambda_1 \in [0, 1], \underline{\mu}_1 \in [0, \mu_1)$$

We analyze the solution to this problem in the next section. Though the information structure consists of $(\lambda_1, \bar{\mu}_1, \underline{\mu}_1)$, any two of them fully characterize the strategy because the
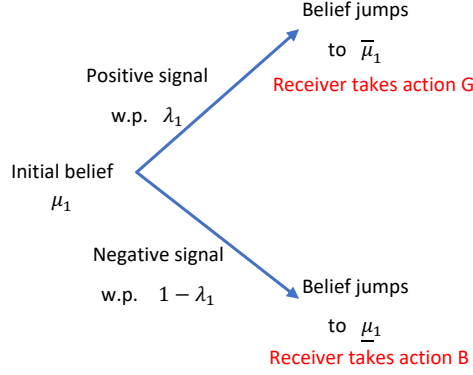
9

Figure 2: Belief Evolution of the Subgame

third variable is then uniquely determined by (F_1). Therefore, we use $(\lambda_1, \bar{\mu}_1)$, the probability of a positive signal and the belief after observing a positive signal, to represent the sender's strategy in the subgame.

2.3.2 Equilibrium of the Full Game

The sender has three options. Firstly, she could provide no information and obtain zero payoffs. Secondly, she could provide information in only one period. If the receiver decides to search, he would observe a one-shot experiment designed by the sender. The receiver takes action G if a positive signal arrives and takes action B if a negative signal comes. The sender would not provide extra information regardless of the experimental outcome. Her problem is exactly (P_1) (except that the subscript 1 shall be replaced by 0). Lastly, the sender could provide information in both periods. She could give a pair of one-shot experiments, which means that the receiver would take action G upon observing a positive signal in either period. If a negative signal arrives in the first period, the sender will provide another experiment, hoping that a positive signal will come in the second period. The sender could also offer a pair of iterative experiments. The receiver takes action G if and only if he observes positive signals in both periods.[7] We illustrate the belief evolution of the one-shot experiments and iterative experiments in Figure 3 and 4.

---

[7] One can see that the receiver will take action B after observing negative signal in the first period under iterative experiments. If, instead, the receiver searches regardless of the realization of the signal in the first period, the sender can save the persuasion cost by providing the second-period experiment in the first period.
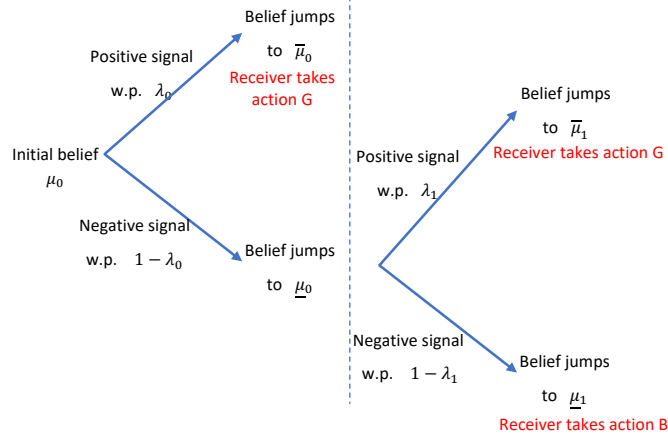
Belief jumps
to $\overline{\mu}_0$

Positive signal

Receiver takes
w.p. λ_0 action G

Belief jumps
to $\overline{\mu}_1$

Positive signal Receiver takes action G

Initial belief
μ_0 w.p. λ_1

Negative signal

w.p. $1-\lambda_0$ Belief jumps

to $\underline{\mu}_0$

Negative signal

w.p. $1-\lambda_1$ Belief jumps

to $\underline{\mu}_1$

Receiver takes action B

Figure 3: Belief Evolution of the One-shot Experiments

Belief jumps
to $\overline{\mu}_1$
Receiver takes action G

Belief jumps

to $\overline{\mu}_0$ Positive signal

Positive signal w.p. λ_1

w.p. λ_0 Negative signal

Initial belief w.p. $1-\lambda_1$
μ_0

Negative signal

w.p. $1-\lambda_0$ Belief jumps Belief jumps

to $\underline{\mu}_0$ to $\underline{\mu}_1$
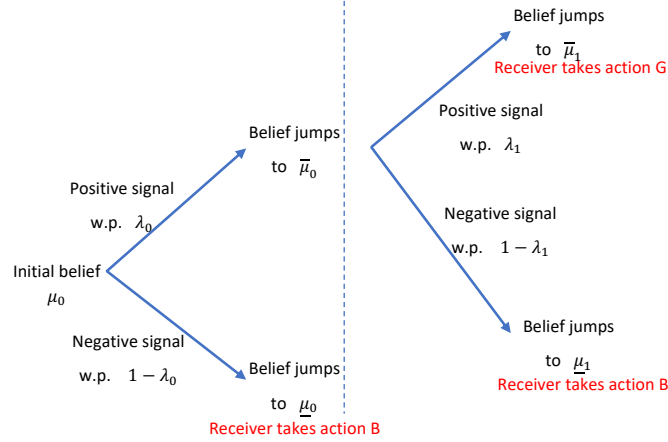
Receiver takes action B Receiver takes action B

Figure 4: Belief Evolution of the Iterative Experiments

The receiver must search in both periods before taking action G under iterative experiments. Hence, compared to one-shot experiments, iterative experiments require a longer search time. To compensate the receiver for the higher expected search costs, the sender needs to provide information more favorable to the receiver, which hurts the sender's payoff. The following result shows that the sender always prefers one-shot experiments in equilibrium. So, we limit our attention to the optimal one-shot experiments.

Proposition 1. *For any pair of feasible iterative experiments, there exists a one-shot experiment that gives the sender a strictly higher payoff.*

The sender takes into account the following constraints when designing the optimal one-shot experiment of the first period:

(1) participation constraint:

$$\lambda_0[\bar{\mu}_0 v_g + (1-\bar{\mu}_0)v_b] + (1-\lambda_0)\mathbb{E}[\text{receiver surplus at } t=1|\text{search at } t=1]$$
$$=\lambda_0(\bar{\mu}_0 + v_b) + (1-\lambda_0)[\lambda_1(\bar{\mu}_1 + v_b) - c] \geq c \qquad (IR_0)$$

(2) feasibility constraint:

$$\lambda_0\bar{\mu}_0 + (1-\lambda_0)\underline{\mu}_0 = \mu_0 \qquad (F_0)$$

If the sender provides information in both periods, her problem is:[8]

$$\max_{\lambda_0,\bar{\mu}_0,\mu_1,\lambda_1,\bar{\mu}_1} -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\lambda_1) + p\lambda_1\right] \qquad (P_2)$$
$$\text{s.t. } (IR_0), (F_0), (\lambda_1,\bar{\mu}_1) \text{ solves } (P_1)$$

We analyze the solution to this problem in the next section. We use $(\lambda_0, \bar{\mu}_0, \mu_1, \lambda_1, \bar{\mu}_1)$, the probability of a positive signal at each period, the belief after observing a positive signal at each period, and the initial belief in the second period, to represent the sender's strategy for the entire game.

3 Optimal Strategies

From the previous discussion, the receiver would search (take action S) whenever the sender provides information. The receiver would take action G immediately upon receiving a positive signal and action B if the sender does not provide information. Therefore, we only need to characterize the sender's strategy, which implies the receiver's strategy.

3.1 A Benchmark

We first consider a benchmark problem in which the receiver always participates and in which the sender can generate an arbitrary amount of information in each period. The sender chooses the information structure to maximize the expected payoff. We will use the

---

[8] We impose the following implicit assumptions on all of the programs in the remaining of the paper: $\bar{\mu}_t \in [-v_b, 1], \underline{\mu}_t \in [0, \mu_t), \lambda_t \in [0, 1], \underline{\mu}_0 = \mu_1$. The last equality comes from the fact that the belief at the beginning of the second period, μ_1, is the belief after observing a negative signal in the first period, $\underline{\mu}_0$, under one-shot experiments.

solution throughout the subsequent analyses.

$$\max_{\lambda_0, \lambda_1} -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0) \left[-K(\lambda_1) + p\lambda_1 \right] \qquad (P_b)$$

Lemma 1. *The solution to the benchmark problem* (P_b), $(\lambda_0^{**}, \lambda_1^{**})$, *does not depend on the search cost* c *and* $\lambda_0^{**} < \lambda_1^{**}$. *The benchmark sender's payoff is strictly positive.*

The benchmark sender's payoff is the highest possible payoff the sender could obtain in equilibrium. When the prior is high enough, a positive signal arrives at the rate of λ_t^{**} and the receiver searches. The sender obtains the benchmark payoff. When the prior is lower and a positive signal arrives at the benchmark probability, the receiver may be quite uncertain about the state even after observing a positive signal (low $\bar{\mu}_t$) due to feasibility constraints. So, the receiver chooses not to search. This friction restricts communication between the sender and the receiver and distorts the optimal strategy away from the benchmark strategy. For the problem to be non-degenerate, we concentrate on the case in which the prior is not too high throughout the paper. As a result, the optimal strategy is different from the benchmark solution.

3.2 Optimal Strategy for the Subgame

When the belief at the beginning of the second period is too low, the receiver will not search, given any feasible experiments. Thus, the sender does not provide information to minimize the cost. When the belief at the beginning of the second period is higher, and the search cost is not too high, the sender provides information and obtains a positive payoff. The following proposition summarizes the optimal information structure.

Proposition 2. *At the second period, the sender does not provide information when* $\mu_1 < \mu_{0,1} := c/v_g$. *When* $\mu_1 \geq \mu_{0,1}$, *the optimal probability of a positive signal and the optimal belief after observing a positive signal,* $(\lambda_1^*, \bar{\mu}_1^*)$, *depend on the search cost* c:

1. *If* $c \geq v_g \lambda_1^{**}$, *there exists a unique* $\widehat{c} \in (v_g \lambda_1^{**}, v_g \mu_1]$ *such that the sender does not provide information if* $c > \widehat{c}$ *and* $(\lambda_1^*, \bar{\mu}_1^*) = (c/v_g, 1)$ *if* $c < \widehat{c}$. *The receiver gets zero surplus.*

2. *If* $c \in [\mu_1 + v_b \lambda_1^{**}, v_g \lambda_1^{**})$, $(\lambda_1^*, \bar{\mu}_1^*) = (\frac{\mu_1 - c}{-v_b}, \frac{-v_b \mu_1}{\mu_1 - c})$. *The receiver gets zero surplus.*

3. *If* $c < \mu_1 + v_b \lambda_1^{**} \wedge v_g \lambda_1^{**}$, $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}} \wedge 1)$. *The receiver gets strictly positive surplus.*

13

When the search cost is too high, the sender has to provide a lot of information to persuade the receiver to search. Even if it is feasible for the sender to provide enough information that the receiver would search, it is so costly that the expected sender's payoff is negative. So, the sender chooses not to provide information, and the receiver does not search.

When the search cost is high but not too high, the sender will provide just enough information such that the receiver searches. Since the receiver's participation constraint is hard to satisfy, in equilibrium, the receiver becomes certain ($\bar{\mu}_1 = 1$) that the state is g after observing a positive signal. Suppose, instead, a positive signal does not fully reveal the state ($\bar{\mu}_1 < 1$). In that case, its arrival rate would need to be higher to persuade the receiver to search. Since the marginal cost of increasing the probability of a positive signal exceeds the marginal benefit, the sender's payoff decreases. The sender trades off the frequency of positive signal for precision.

When the search cost is intermediate, the receiver's participation constraint is easier to satisfy. Since the marginal benefit of increasing the probability of a positive signal exceeds the marginal cost, in equilibrium, the sender trades off the precision of a positive signal for frequency. The receiver is still uncertain about the state after observing a positive signal, but the belief is high enough that the receiver searches.

When the search cost is low, the information friction does not distort the information structure. The sender provides the benchmark amount of information, and the receiver gets a strictly positive surplus.

3.3 Optimal Strategy for the Entire Game

When the prior is too low, any feasible signal the sender could generate is not attractive enough for the receiver to search. Thus, it is impossible to communicate between the sender and the receiver. When the prior is higher, and the search cost is not too high, the sender provides information and obtains a positive payoff.

Proposition 3. *Suppose the search cost is not too high, $c < \widehat{c}$. There exists $\mu_{1,2} \geq c(2v_g - c)/(v_g)^2$ such that the sender does not provide information if the prior is low, $\mu_0 < \mu_{0,1}$, provides information in one period if $\mu_0 \in [\mu_{0,1}, \mu_{1,2})$, and provides information in both periods if $\mu_0 > \mu_{1,2}$. Suppose the sender provides information in both periods. A positive signal fully reveals the state (which is g) when the search cost is high and partially reveals the state when the search cost is low. The receiver gets zero total surplus.*

Since the information provision cost is non-linear in the probability of a positive signal, the sender has an incentive to smooth the information provision over two periods. When the

14

prior is low, it is not feasible for the sender to provide enough information in both periods so that the receiver would search whenever a positive signal has not arrived. As the prior increases, it becomes feasible for the sender to smooth the information provision. If the sender finds it optimal to provide information in both periods at a given prior, she also prefers to smooth the information for any higher prior.

When it is highly costly for the receiver to search, the optimal information structure fully convinces the receiver that the state is g when a positive signal arrives. In equilibrium, the receiver obtains the highest possible surplus conditional on observing a positive signal and making the purchase. Without providing this type of information, which is favorable to the receiver, the sender cannot persuade the receiver to search. In contrast, when it is less costly for the receiver to search, the optimal information structure adds some noise to a positive signal. In equilibrium, the receiver is not sure that the state is g after observing a positive signal. The state may be b after the receiver takes action G. However, the likelihood of state g after a positive signal is high enough to persuade the receiver to search. By adding some noise to a positive signal, the sender can provide more frequent positive signals and increase her payoff without violating the feasibility constraint.

When the sender provides information in both periods, she can always extract surplus from the receiver if the receiver gets a strictly positive surplus. If the sender faces information under-provision, she can increase the payoff by increasing the probability of a positive signal and decreasing the belief after observing a positive signal. If the sender faces information over-provision, she can increase the payoff by reducing the probability of a positive signal and increasing the belief after observing a positive signal. This implies that the receiver gets zero surplus in equilibrium.

The optimal strategy is consistent with real-world examples. Consider the application of product sales in section 2.2.1 when a buyer decides whether to purchase a pair of shoes and searches on Google to gather some information. The search cost is relatively low, and the buyer is unlikely to be sure that it is a good match, even after obtaining a positive signal and making the purchase. On the contrary, when a buyer visits a dealership and acquires information on the car of interest through talking to the dealer and test driving it, the search cost is relatively high, because it is costly to travel to the physical store and spend a significant amount of time there. When the buyer decides to buy the car, he is usually confident that it is a precisely good match.

3.4 Comparative Statics

When the sender provides information in only one period, the optimal strategy has a closed-form solution and is easy to analyze. Here, we discuss the comparative statics when the sender provides information in both periods. The specific form of the optimal strategy depends on the relative size of the search cost. According to Proposition 2, the sender's strategy in the subgame is constant when the search cost is high (but not too high). When the search cost is lower, the sender's strategy depends on her belief at the beginning of the second period, μ_1. When $\mu_1 > c - v_b \lambda_1^{**}$, the receiver expects to get a strictly positive surplus in the second period, and thus the first-period participation constraint is relaxed (denote the corresponding strategy as the S_+ strategy). When $\mu_1 \in [c/v_g, c - v_b \lambda_1^{**}]$, the receiver expects to get zero surplus in the second period (denote the corresponding strategy as the S_0 strategy). In equilibrium, the sender endogenously determines whether to use the S_0 or S_+ strategy.

3.4.1 Comparative Statics With Regard to the Prior

When the search cost is high, the prior determines whether the sender smooths information but does not affect the information structure, conditional on the sender providing information. When the search cost is low, the prior affects the information structure monotonically. When the search cost is intermediate, the sender may switch from the S_0 strategy to the S_+ strategy as the prior increases. There could be a discrete jump in the optimal information structure. We leave the analysis of this case to the appendix.

Proposition 4. *Suppose the sender provides information in both periods. When the search cost is high, $v_g \lambda_1^{**} \leq c < \widehat{c}$, positive signal fully reveals the state. Neither the probability of a positive signal, λ_t^*, nor the sender's payoff depends on the prior, $(\lambda_t^*, \bar{\mu}_t^*) = (c/v_g, 1)$. When the search cost is low, $c \leq \tilde{c} := v_g K'^{-1} [K(\lambda_1^{**})/\lambda_1^{**}]$, the probability of a positive signal, λ_t^*, is continuous and increases in the prior. The belief after observing a positive signal, $\bar{\mu}_t^*$, is continuous and decreases in the prior. The sender's payoff strictly increases in the prior.*

The optimal information is perfectly smooth when the search cost is high and the sender provides information in both periods. Since the participation constraint of the receiver is strong, a positive signal fully reveals the state is g. Because it is very costly for the receiver to acquire information, the minimal amount of information to persuade the receiver to search is high. The marginal cost of providing more information exceeds the marginal benefit. Even if the prior increases and it is feasible for the sender to provide more information, she would prefer not to do so. Hence, conditional on the sender providing information, the information structure does not depend on the prior.

When the search cost is low, and the sender provides information in both periods, she chooses between S_+ and S_0 strategies. Under the S_+ strategy, the receiver observes less frequent positive signals in the first period and more frequent positive signals in the second period. On average, he spends a longer time searching. Consequently, the sender has to provide information more favorable to the receiver to compensate for the higher expected total search cost, which reduces the sender's surplus. Therefore, the sender always chooses the S_0 strategy in equilibrium, and the optimal strategy is continuous in the prior. The sender faces information under-provision in both periods. More frequent positive signals are feasible when the prior is higher. Even if the receiver becomes less sure about the state being good after observing a positive signal, he would still search as long as the likelihood of receiving a positive signal and earning a strictly positive surplus increases. In equilibrium, the sender trades off the precision of a positive signal for frequency as the prior increases.

3.4.2 Comparative Statics With Regard to Sender's Costs

Providing the same amount of information may impose different costs on the sender. To study the impact of the sender's information provision costs on the optimal strategy, we rewrite the sender's cost function as $K(\lambda) = \eta \tilde{K}(\lambda)$, where $\tilde{K}(1/2) = 1$ for identification. It is more costly for the sender to provide information when η is larger. The following proposition summarizes the comparative statics of the optimal strategy about η.

Proposition 5. *Suppose the sender provides information in both periods. Her payoff strictly decreases in the relative cost of information provision, η. When the search cost is low, $c \leq \tilde{c}$, the sender provides (weakly) less information in the first period and (weakly) more information in the second period, as η increases. When the search cost is high, $c \geq v_g \lambda_1^{**}$, the optimal strategy of the sender does not depend on η.*

When the search cost is high, it is very costly for the receiver to search. The sender needs to provide a lot of information to persuade the receiver to search. As a result, the marginal cost of providing more information exceeds the marginal benefit. So, the sender provides the minimum amount of information for the receiver to search, which does not depend on the sender's cost. Hence, the optimal information structure does not depend on the relative cost of information provision.

When the search cost is low, and the relative cost of information provision increases, the marginal cost of providing information increases, while the marginal benefit remains the same. Because the sender definitely incurs the information provision cost in the first period, she provides less information in the first period. This allows her to provide more information

in the second period when the information provision cost is not always incurred, and when she faces information under-provision.

3.5   A Numerical Example

We illustrate the optimal strategy of the sender by a numerical example. The sender's cost function has the truncated quadratic form, $K(\lambda) = k\lambda^2/(1 - \lambda)$, which satisfies Assumption 1. In each figure below, we present the optimal strategy and the sender's payoffs from the optimal one-period, S_0, and S_+ strategies[9]. The search cost is low in Figure 5.[10] The sender always prefers S_0 strategy when she provides information in both periods. As illustrated, the probabilities of positive signal at both periods, λ_0^* and λ_1^*, are continuous and increase in μ_0. The beliefs after observing a positive signal in each period, $\bar{\mu}_0^*$ and $\bar{\mu}_1^*$, are continuous and decrease in μ_0. When the prior is lower than the intercept of the brown line, the sender prefers providing information in only one period to smooth the information provision. When the search cost increases to the intermediate level (Figure 6), the sender always provides information in both periods provided that it is feasible. So, we do not plot the sender's payoff of providing information in only one period. The sender prefers S_0 strategy for a low prior and switches to S_+ strategy as the prior increases. The optimal strategy is non-monotonic and discontinuous in μ_0 due to the switch. When the search cost further increases (Figure 7), the optimal information structure is perfectly smooth, and a positive signal always fully reveals the state.
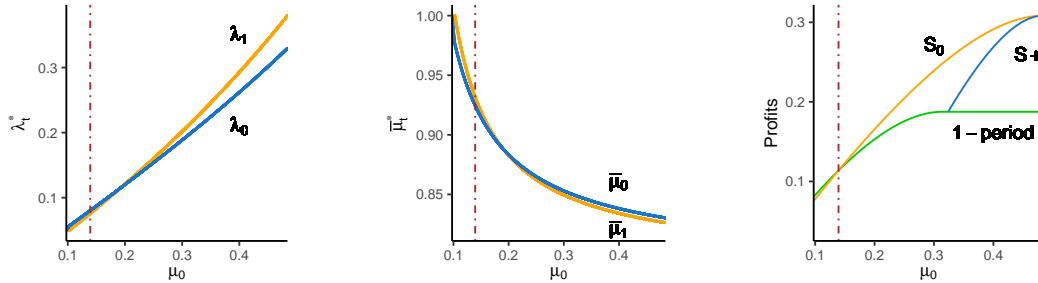

Figure 5: The optimal strategy when $c = 0.01, p = 0.8, v_g = 0.2, v_b = -0.8, k = 0.5$

---

[9] The domain of the prior is $[c(2v_g - c)/(v_g)^2, \widehat{\mu}_0 \wedge p]$. When $\mu_0 < c(2v_g - c)/(v_g)^2$, the sender provides information in at most one period. When $\mu_0 \geq \widehat{\mu}_0 := 2c - v_b\lambda_1^{**} - [c + (1 - \lambda_1^{**})v_b]\lambda_0^{**}$, the sender always chooses the benchmark solution.

[10] The choice of the specific parameter values do not affect the qualitative property of the optimal strategy (i.e., the shape of the figure). What matters is the relative value. The search cost is low if $c \leq (1 - p)\lambda_0^{**}$, intermediate if $(1 - p)\lambda_0^{**} < c < (1 - p)\lambda_1^{**}$, and high if $(1 - p)\lambda_1^{**} \leq c < \widehat{c}$.
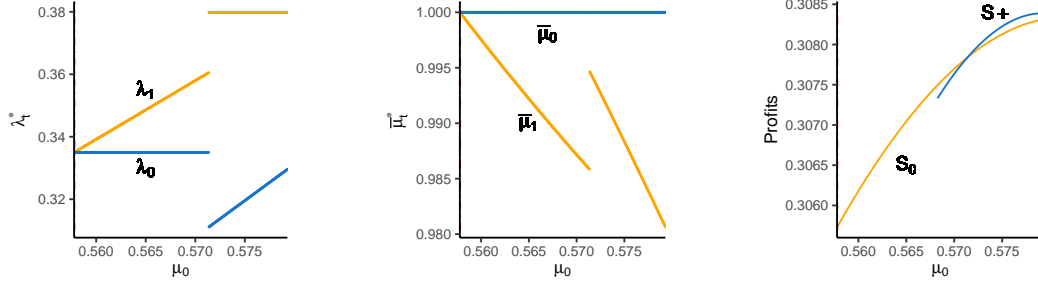
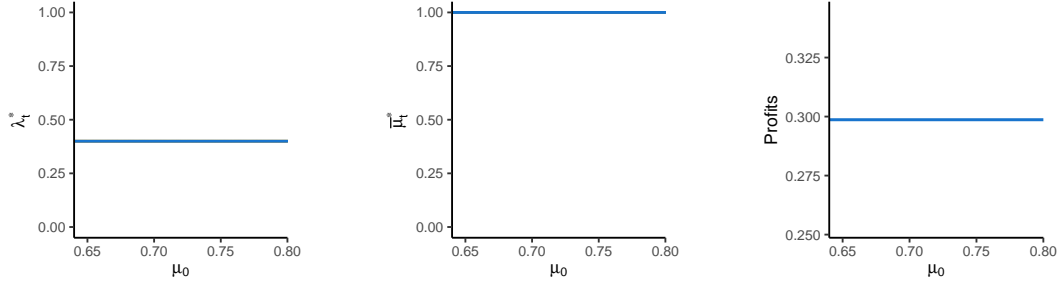Figure 6: The optimal strategy when $c = 0.067, p = 0.8, v_g = 0.2, v_b = -0.8, k = 0.5$

Figure 7: The optimal strategy when $c = 0.08, p = 0.8, v_g = 0.2, v_b = -0.8, k = 0.5$

4   The Efficient Information Structure

In the previous sections, the sender designs the information structure to maximize the expected payoff. This section characterizes the efficient strategy when a social planner designs the information structure to maximize total welfare. We then investigate the information distortion caused by not taking into account receiver surplus. For tractability reasons, we use a special form of the payoff function, specified in section 2.2.1. So, in this section, $v_g = 1 - p$ and $v_b = -p$.

4.1   Efficient Strategy for the Subgame

As discussed in the previous section, the sender does not provide information if the belief at the beginning of the second period is too low, $\mu_1 < \mu_{0,1}$. So, we concentrate on the case

in which $\mu_1 \geq \mu_{0,1}$. In the second period, the social planner's problem is:

$$\max_{\lambda_1,\bar{\mu}_1} -K(\lambda_1) + p\lambda_1 + \lambda_1(\bar{\mu}_1 - p) - c \qquad (E_1)$$

$$\text{s.t. } (IR_1), (F_1)$$

Below, we discuss the efficient information structure of the subgame intuitively. The formal characterization of the efficient strategy of the subgame is in the appendix. When the search cost is high, it is very costly for the receiver to search. The sender needs to provide a lot of information to persuade the receiver to search. As a result, the marginal cost of providing more information exceeds the marginal benefit. So, the sender provides the minimum amount of information necessary to induce the receiver to search, which does not depend on whether the sender maximizes the sender's surplus or total welfare. Hence, there is no information distortion. When the search cost is lower, it is easier to persuade the receiver to search. The sender provides more than the minimum amount of information necessary to induce the receiver to search. The marginal costs of providing more information are the same for both the sender and the social planner, while the marginal benefit of providing more information is smaller for the sender. Therefore, the sender provides less information than the social planner does when the search cost is high or the initial belief is high. One exception is that, when both the search cost and the initial belief are low, the sender provides more information than the social planner does because the social planner can only generate infrequent signals.

4.2 Efficient Strategy for the Entire Game

As in the previous section, the social planner does not provide information if the prior is too low or the search cost is too high. When she provides information in only one period, the previous subsection (and Proposition 15 in the appendix, where the subscript 1 needs to be replaced by 0) characterizes the efficient strategy. When she provides information in both periods, her problem is:

$$\max_{\lambda_0,\bar{\mu}_0,\mu_1,\lambda_1,\bar{\mu}_1} -K(\lambda_0) + \lambda_0\bar{\mu}_0 - c + (1-\lambda_0)\left[-K(\lambda_1) + \lambda_1 - c\right] \qquad (E_2)$$

$$\text{s.t. } (IR_0), (F_0), (\lambda_1, \bar{\mu}_1) \text{ solves } (E_1)$$

The following proposition compares the payoff-maximizing strategy and the efficient one when the sender provides information in both periods.

Proposition 6. *Suppose the sender provides information in both periods. When $c \geq v_g\lambda_1^{**}$, the payoff-maximizing strategy is efficient. When $c < v_g\lambda_1^{**}$, the sender, who maximizes*

the payoff, provides less information in the first period and more information in the second period than does the social planner, who maximizes total welfare.

When the search cost is high, similar to the argument in the previous subsection, the sender provides the minimum amount of information for the receiver to search, which does not depend on the sender's objective (maximizing the sender surplus or total welfare). Hence, there is no information distortion. When the search cost is lower, it is easier to persuade the receiver to search. The sender provides more than the minimum amount of information for the receiver to search. The marginal costs of providing more information are the same for both the sender and the social planner, while the marginal benefit of providing more information is smaller for the sender. Therefore, the sender provides less information than the social planner does when the search cost is high or when the initial belief is high. One exception is that, when both the search cost and the initial belief are low, the sender provides more information than the social planner does because the social planner can only generate infrequent signals.

5   Model Extensions

5.1   Dynamic Commitment

Under many circumstances, the assumption that the sender can generate credible signals within each period but does not have dynamic commitment power is reasonable. It is hard for the sender to commit to the entire information structure across all periods ex-ante and to convince the receiver that the sender will stick to the information structure when it is profitable to deviate to a different information structure during intermediate periods. However, factors such as reputation could give the sender stronger commitment power. Here, we study the implications of the case in which the sender has dynamic commitment power. The sender chooses and commits to the entire information structure to maximize the ex-ante expected surplus.

$$\max_{\lambda_0,\bar{\mu}_0,\mu_1,\lambda_1,\bar{\mu}_1} -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\lambda_1) + p\lambda_1\right] \qquad (P_{dc})$$

$$\text{s.t. } (IR_0), (F_0), (IR_1), (F_1)$$

Proposition 7. *Suppose the search cost is high, $v_g\lambda_1^{**} < c < \widehat{c}$, and $\mu_0 > c(2v_g - c)/(v_g)^2$. The sender provides information in both periods regardless of the dynamic commitment power. If the sender has dynamic commitment power, her payoff is strictly higher, and*

the receiver gets a strictly positive surplus in the second period. The benefit of dynamic commitment power for the sender vanishes as the search cost approaches zero.

We have shown that, when the search cost is high, the sender perfectly smooths the information if she doesn't have dynamic commitment power. If the sender instead has dynamic commitment power, she would commit to providing information more favorable to the receiver in the second period. As a result, the receiver will search even if the sender provides less information in the first period, relaxing the information over-provision issue. Though it hurts the sender's payoff in the second period, it increases the sender's payoff in the first period by reducing the information provision cost. The overall effect is strictly positive. So, the optimal information provision will not be perfectly smooth.

The above finding is related to results on durable good pricing (e.g., Coase 1972). Without dynamic commitment power, the monopolist tends to reduce the price as time goes on, which reduces profit, because a rational receiver will strategically wait. Here, dynamic commitment power also benefits the sender, but the underlying mechanisms differ. This paper focuses on persuasion (information provision) rather than incentive (pricing). In addition, in the durable good pricing example, the ability to commit not to provide a more favorable price in the future benefits the sender. By contrast, in this paper, the ability to commit to more favorable information in the future benefits the sender.

However, when the search cost approaches zero, the difference between the sender surplus with and without dynamic commitment power approaches zero. The intuition is that the benefit of dynamic commitment power comes from relaxing the participation constraint in the first period by committing to providing more favorable information in the second period. However, the participation constraint is already very loose when the search cost is low. Thus, the benefit of dynamic commitment power approaches zero.

5.2   Discounting

In the main model, we assume that there is no discounting. In reality, information acquisition and provision usually happen in a short period, and that assumption is reasonable. However, some communications between the sender and the receiver could take longer. Here, we study the information provision strategy when the sender and the receiver have the same

discount factor, $\delta \in (0, 1)$. With discounting, the sender's problem becomes:

$$\max_{\lambda_0, \bar{\mu}_0, \mu_1, \lambda_1, \bar{\mu}_1} -K(\lambda_0) + p\lambda_0 + \delta(1 - \lambda_0) \left[-K(\lambda_1) + p\lambda_1 \right] \qquad (P_{2,\delta})$$

$$\text{s.t. } \lambda_0(\bar{\mu}_0 + v_b) + \delta(1 - \lambda_0)[\lambda_1(\bar{\mu}_1 + v_b) - c] \geq c \qquad (IR_{0,\delta})$$

$$(F_0), (\lambda_1, \bar{\mu}_1) \text{ solves } (P_1)$$

One can see that both the objective function and the first-period participation constraint change. When the players become less patient (the discount factor δ decreases), the present value of the second-period sender surplus decreases, and the first-period participation constraint becomes tighter. Thus, it is less attractive for the sender to sell the goods in the second period.

Proposition 8. *When the search cost is high, $c \geq v_g \lambda_1^{**}$, the optimal strategy of the sender does not depend on the discount factor, δ. When the search cost is low, $c \leq \tilde{c}$, the sender provides (weakly) more information in the first period and (weakly) less information in the second period, as the players become less patient.*

When the search cost is high, it is hard to satisfy the participation constraints of the receiver. Providing enough information to persuade the receiver to search dominates the force of discounting. Therefore, the sender's strategy remains the same as the no-discounting case, and the sender perfectly smooths information provision. When the search cost is low, the sender provides (weakly) less information in the second period when the players are less patient because of discounting. The sender provides more information in the first period as she becomes more tempted to convert the receiver early.

We assume in the main model that the receiver is forward-looking and takes into account the potential payoff of the second period when he chooses his action in the first period. This assumption is reasonable if the information environment is transparent and the receiver knows that gradual learning is possible. Here, we consider the possibility of asymmetric discount factors. In particular, the sender is perfectly patient while the receiver is perfectly impatient (myopic). If the receiver is myopic, then he trades off only the current-period benefit and cost in deciding whether or not to search. The information structure in the second period cannot relax the first-period participation constraint. Therefore, when the receiver is myopic, the feasible information structure is a subset of when the receiver is forward-looking. If the receiver surplus in the second period is strictly positive when the receiver is forward-looking, the optimal information structure may not be feasible when the receiver is myopic. Hence, the sender is (weakly) worse off if the receiver is myopic rather than forward-looking. This result has managerial implications, as it suggests that the sender

should try to better inform the receiver of the possibility of gradual learning. Common knowledge of gradual information revelation improves the sender surplus.

5.3 Infinite Number of Periods

The two-period model can capture information smoothing and gradual learning. We extend the main model to a model with an infinite number of periods and show that the main insights extend to this richer model. This also provides some additional insights.

Time is discrete, $t = 0, 1, 2, ...$ In each period, the sender determines and commits to the information structure of the experiment but cannot commit to the experiment in the future. The receiver can search for information before deciding. Unlike the two-period model, there is no deadline. The receiver can search for as long as he wants. Since the sender's payoff is bounded by p, the payoff function is well-defined even without the discount factor. So, we do not consider discounting for consistency with the main model. We can analyze the problem similarly if we include discounting.

Proposition 9. *The receiver searches for at most $\lfloor \frac{v_g \mu_0}{c} \rfloor$ periods. When the search cost is high, $v_g \lambda_1^{**} \leq c < \widehat{c}$, the sender provides perfectly smooth information for $k := \lfloor \frac{ln(1-\mu_0)}{ln(1-c/v_g)} \rfloor$ periods, and a positive signal fully reveals the state, $(\lambda_t^*, \bar{\mu}_t^*) = (c/v_g, 1)$, for $t = 0, 1, ..., k-1$. When the search cost is low, a positive signal partially reveals the state. As the search cost approaches zero, the sender could obtain the equilibrium payoff as if the persuasion cost were zero.*

This proposition shows that the main insights are robust to the specification of the length of time. The two-period model corresponds to the case when there is a deadline in the information acquisition. When time is infinite, there is no limit on how long the receiver can search. Under high search costs, the optimal information structure fully convinces the receiver that the state is g when a positive signal arrives. Under low search costs, the optimal information structure adds some noise to the positive signal. The same intuition as for proposition 3 applies. Because the receiver can keep searching for a longer period, the sender can better smooth the information. When the search cost is high, the sender may provide information for more than two periods if she believes that the state is likely to be g and the receiver is willing to spend more time searching. When the search cost approaches zero, the ability to smooth the information is valuable for the sender. She could obtain the equilibrium payoff as if the persuasion cost were zero. Because of the low search cost, the sender can convince the receiver to search with very little information in each period. The sender's cost becomes very low by smoothing the information over many periods.

24

The model predicts that the receiver will decide within a bounded number of periods under endogenous information structures. On the contrary, consumers may be stuck in the continuation region and keep searching when the information environment is exogenous (e.g., Moscarini and Smith 2001, Branco et al. 2012). The consumer incurs a high search cost when he spends a long time searching. Too much searching hurts not only the consumer but also the firm. To persuade the consumer to keep searching, the firm must incur higher persuasion costs. In equilibrium, the firm will design the information structure such that the consumer searches for a limited period.

6   Concluding Remarks

This paper examines the optimal information provision strategy of a sender and the optimal information acquisition strategy of a receiver when the sender sequentially persuades a receiver to take a particular action, i.e., to purchase a good. The sender prefers that action regardless of the unknown state, while the receiver only wishes to take the action if the state is good. In our model, the sender incurs a cost to provide information, and the receiver incurs a cost to search. The receiver trades off the cost of searching and the benefit of obtaining more accurate information to make better decisions. The sender trades off the cost of information provision and the benefit of persuading the receiver to search and then take the desired action. We allow for sequential information provision and acquisition, so the communication between the receiver and the sender is gradual. Consequently, the sender also makes the intertemporal tradeoff of smoothing the information to reduce the total cost. In equilibrium, she uses one-shot experiments that induce the receiver to take the desired action immediately upon observing a positive signal. The sender smooths information over multiple periods if and only if there is a high prior that the state is good. The sender extracts all the surplus from the receiver when she provides information in both periods, while she may leave some rent for the receiver when she provides information in only one period. When the search cost of the receiver is high, the receiver is sure that the state is good when he takes the desired action. When the search cost is low, the optimal information structure does not fully reveal the state, which may be bad even though the receiver takes the desired action. We compare the payoff-maximizing information structure with the efficient information structure and find no information distortion when the search cost is high. There could be upward or downward information distortion when the search cost is lower.

This paper endogenizes the consumer's information environment from the firm's perspective. Instead of looking at specific parameters of the search environment, we study the design of the information environment under a general signal space. We characterize the optimal

information structure among all feasible information policies. The implementation of a given signal depends on the institutional details of the specific problem. Further empirical work can complement the current paper by putting the theoretical results into practice. In addition, it would be interesting to study the optimal persuasion strategy when there is more than one sender. Such competition may lead the sender to provide more information and thereby improve equilibrium efficiency. Moreover, the sender has complete control of the information structure in this paper. It would be interesting to consider the case when the sender can only partially control the information environment.

References

Aumann, R. J., Maschler, M. (1995). Repeated games with incomplete information. MIT press.

Bergemann, D., & Morris, S. (2019). Information design: A unified perspective. Journal of Economic Literature, 57(1), 44-95.

Branco, F., Sun, M., Villas-Boas, J. M. (2012). Optimal search for product information. Management Science, 58(11), 2037-2056.

Branco, F., Sun, M., Villas-Boas, J. M. (2016). Too much information? Information provision and search costs. Marketing Science, 35(4), 605-618.

Che, Y. K., Kim, K., Mierendorff, K. (2020). Keeping the listener engaged: a dynamic model of bayesian persuasion. arXiv preprint arXiv:2003.07338.

Chen, Y., & Yao, S. (2017). Sequential search with refinement: Model and application with click-stream data. Management Science, 63(12), 4345-4365.

Coase, R. H. (1972). Durability and monopoly. The Journal of Law and Economics, 15(1), 143-149.

Dukes, A., & Liu, L. (2016). Online shopping intermediaries: The strategic design of search environments. Management Science, 62(4), 1064-1077.

Ely, J. C. (2017). Beeps. American Economic Review, 107(1), 31-53.

Ely, J. C., Szydlowski, M. (2020). Moving the goalposts. Journal of Political Economy, 128(2), 468-506.

Gardete, P., & Hunter, M. (2020). Guiding Consumers through Lemons and Peaches: An Analysis of the Effects of Search Design Activities.

Gentzkow, M., Kamenica, E. (2014). Costly persuasion. American Economic Review, 104(5), 457-62.

Guo, L. (2021). Endogenous Evaluation and Sequential Search. Marketing Science, 40(3), 413-427.

Hong, H., & Shum, M. (2006). Using price distributions to estimate search costs. The RAND Journal of Economics, 37(2), 257-275.

Honka, E. (2014). Quantifying search and switching costs in the US auto insurance industry. The RAND Journal of Economics, 45(4), 847-884.

Honka, E., & Chintagunta, P. (2017). Simultaneous or sequential? search strategies in the us auto insurance industry. Marketing Science, 36(1), 21-42.

Iyer, G., Zhong, Z. Z. (2021). Pushing Notifications as Dynamic Information Design. Marketing Science

Kamenica, E., Gentzkow, M. (2011). Bayesian persuasion. American Economic Review,

101(6), 2590-2615.

Ke, T. T., Villas-Boas, J. M. (2019). Optimal learning before choice. Journal of Economic Theory, 180, 383-437.

Ke, T. T., Shen, Z. J. M., Villas-Boas, J. M. (2016). Search for information on multiple products. Management Science, 62(12), 3576-3603.

Kim, J. B., Albuquerque, P., & Bronnenberg, B. J. (2010). Online demand under limited consumer search. Marketing science, 29(6), 1001-1023.

Kim, J. B., Albuquerque, P., & Bronnenberg, B. J. (2017). The probit choice model under sequential search with an application to online retailing. Management Science, 63(11), 3911-3929.

Kuksov, D., Villas-Boas, J. M. (2010). When more alternatives lead to less choice. Marketing Science, 29(3), 507-524.

Ma, L. (2016). Only the interested learn: A model of proactive learning of product reviews. Working paper.

Moraga-González, J. L., Sándor, Z., & Wildenbeest, M. R. (2021). Consumer search and prices in the automobile market.

Moscarini, G., Smith, L. (2001). The optimal level of experimentation. Econometrica, 69(6), 1629-1644.

Orlov, D., Skrzypacz, A., Zryumov, P. (2020). Persuading the principal to wait. Journal of Political Economy, 128(7), 2542-2578.

Seiler, S. (2013). The impact of search costs on consumer behavior: A dynamic approach. Quantitative Marketing and Economics, 11(2), 155-203.

Seiler, S., & Pinna, F. (2017). Estimating search benefits from path-tracking data: measurement and determinants. Marketing Science, 36(4), 565-589.

Stigler, G. J. (1961). The economics of information. Journal of political economy, 69(3), 213-225.

Ursu, R. M., Wang, Q., & Chintagunta, P. K. (2020). Search duration. Marketing Science, 39(5), 849-871.

Villas-Boas, J. M., & Yao, Y. (2021). A dynamic model of optimal retargeting. Marketing Science, 40(3), 428-458.

Wei, D. (2021). Persuasion under costly learning. Journal of Mathematical Economics, 94, 102451.

Weitzman, M. L. (1979). Optimal search for the best alternative. Econometrica: Journal of the Econometric Society, 641-654.

Zhong, W. (2019). Optimal dynamic information acquisition. Available at SSRN.

Zhong, Z. Z. (2022). Platform Search Design: The Roles of Precision and Price, Marketing Science, forthcoming

Appendix

The following proposition formalizes the claim that we could work with mean-preserving posterior beliefs rather than the specific signal structure in section 2.1.

Proposition 10. *(Equivalent Representations of the Signal) The following are equivalent:*

1. *There exists a binary signal $s \in \Delta(\{0,1\})$ such that $\mathbb{P}[s = 1|g] > \mathbb{P}[s = 1|b]$.*

2. *There exists a binary-valued posterior belief whose expectation is equal to the prior.*

Proof of Proposition 10. $1 \Rightarrow 2$: Given a binary signal such that $\mathbb{P}[s = 1|g] > \mathbb{P}[s = 1|b]$, law of iterated expectation implies that $\mathbb{E}[\mathbb{P}[g|s]] = \mathbb{E}[\mathbb{E}[\mathbf{1}_{[g]}|s]] = \mathbb{E}[\mathbf{1}_{[g]}] = \mathbb{P}[g]$. So, the expectation of the posterior belief is equal to the prior. Note that $\mathbb{P}[s = 0|g] < \mathbb{P}[s = 01|b]$. By Bayes' rule, $\mathbb{P}[g|s = 1] = \frac{\mathbb{P}[s=1|g]\mathbb{P}[g]}{\mathbb{P}[s=1|g]\mathbb{P}[g]+\mathbb{P}[s=1|b]\mathbb{P}[b]} > \frac{\mathbb{P}[s=1|g]\mathbb{P}[g]}{\mathbb{P}[s=1|g]\mathbb{P}[g]+\mathbb{P}[s=1|g]\mathbb{P}[b]} = \mathbb{P}[g] = \frac{\mathbb{P}[s=0|g]\mathbb{P}[g]}{\mathbb{P}[s=0|g]\mathbb{P}[g]+\mathbb{P}[s=0|g]\mathbb{P}[b]} > \frac{\mathbb{P}[s=0|g]\mathbb{P}[g]}{\mathbb{P}[s=0|g]\mathbb{P}[g]+\mathbb{P}[s=0|b]\mathbb{P}[b]} = \mathbb{P}[g|s = 0]$. So, the posterior belief is binary-valued.

$2 \Rightarrow 1$: Given a binary-valued posterior belief whose expectation is equal to the prior, μ_0. Denote the distribution of the belief by $\mu = \begin{cases} \bar{\mu}_1 > \mu_0 & w.p. \ \lambda_1 \\ \underline{\mu}_1 < \mu_0 & w.p. \ 1 - \lambda_1 \end{cases}$. We now construct a binary signal $s \in \Delta(\{0,1\})$. Define $\mathbb{P}[s = 1|g] = \frac{\bar{\mu}_1 \lambda_1}{\mu_0}$ and $\mathbb{P}[s = 1|b] = \frac{(1-\bar{\mu}_1)\lambda_1}{(1-\mu_0)}$. One can verify by Bayes' rule that this signal s induces exactly the same posterior belief, using the assumption that $\mu_0 = \lambda_1 \bar{\mu}_1 + (1 - \lambda_1)\underline{\mu}_1$. We just need to show that $\mathbb{P}[s = 1|g] > \mathbb{P}[s = 1|b]$, which follows from the fact that $\bar{\mu}_1 > \mu_0$. $\square$

Proof of Proposition 1. We first characterize the optimal one-period strategy (providing an one-shot experiment) of the sender. Analagous to section 2.3.1, the sender's problem is:

$$\max_{\lambda_0, \bar{\mu}_0} -K(\lambda_0) + p\lambda_0 \qquad (P_0)$$

$$\text{s.t. } \lambda_0(\bar{\mu}_0 + v_b) \geq c \qquad (IR'_0)$$

$$\lambda_0 \bar{\mu}_0 + (1 - \lambda_0)\underline{\mu}_0 = \mu_0, \qquad (F_0)$$

$$\lambda_0 \in [0, 1], \underline{\mu}_0 \in [0, \mu_0)$$

We transform (P_0) into an equivalent constrained program that is easier to analyze.

Lemma 2. *If $\mu_0 < c/v_g$, the sender does not provide information in the second period. If $\mu_0 \geq c/v_g$, (P_0) is equivalent to:*

$$\Pi_1(\mu_0) := \max -K(\lambda_0) + p\lambda_0 \qquad (P_0')$$

$$s.t. \ \lambda_0 \in \left[\frac{c}{v_g}, \frac{\mu_0 - c}{-v_b}\right]$$

Proof. We first show that any (λ_0, μ_0) satisfying the constraints in (P_0), (IR_0'), (F_0), also satisfy the constraints in (P_0'):

$(IR_0') \Rightarrow \lambda_0 \geq \frac{c}{\bar{\mu}_0 + v_b} \geq \frac{c}{v_g}$. (IR_0') & $(F_0) \Rightarrow \lambda_0 \leq \frac{\mu_0 - c - \underline{\mu}_0}{-v_b - \underline{\mu}_0} \leq \frac{\mu_0 - c}{-v_b}$. Thus, $\lambda_0 \in \left[\frac{c}{v_g}, \frac{\mu_0 - c}{-v_b}\right]$. It is feasible for the sender to provide information in the second period iff $\left[\frac{c}{v_g}, \frac{\mu_0 - c}{-v_b}\right]$ is non-empty: $\frac{c}{v_g} \leq \frac{\mu_0 - c}{-v_b} \Leftrightarrow \mu_0 \geq \frac{c}{v_g}$. So, If $\mu_0 < \frac{c}{v_g}$, the sender would not provide information in the second period.

We then show that for any (λ_0, μ_0) satisfying the constraints in (P_0') and $\mu_0 \geq \frac{c}{v_g}$, we could find $\bar{\mu}_0, \underline{\mu}_0$ such that $(\lambda_0, \mu_0, \bar{\mu}_0, \underline{\mu}_0)$ satisfies the constraints in (P_0). The conclusion then follows.

Suppose (λ_0, μ_0) satisfies the constraints in (P_0'): $\lambda_0 \in \left[\frac{c}{v_g}, \frac{\mu_0 - c}{-v_b}\right], \mu_0 \geq \frac{c}{v_g}$. Consider $\bar{\mu}_0 = \frac{c}{\lambda_0} - v_b$ and $\underline{\mu}_0 = \frac{\mu_0 - c + v_b\lambda_0}{1 - \lambda_0}$. One can verify that $(\lambda_0, \mu_0, \bar{\mu}_0, \underline{\mu}_0)$ satisfies (IR_0') & (F_0). So, we just need to show that $-v_b \leq \bar{\mu}_0 \leq 1$ and $\underline{\mu}_0 \geq 0$. $\bar{\mu}_0 = \frac{c}{\lambda_0} - v_b \geq -v_b$. $\lambda_0 \geq \frac{c}{v_g} \Rightarrow \bar{\mu}_0 = \frac{c}{\lambda_0} - v_b \leq 1$. $\lambda_0 \leq \frac{\mu_0 - c}{-v_b} \Rightarrow \mu_0 \geq c - v_b\lambda_0 \Rightarrow \underline{\mu}_0 = \frac{\mu_0 - c + v_b\lambda_0}{1 - \lambda_0} \geq 0$. $\qquad \square$

Now consider the transformed program (P_0') when $\mu_0 \geq \frac{c}{v_g}$.

1. If $c \geq v_g\lambda_1^{**}$ (i.e. $\lambda_1^{**} \leq \frac{c}{v_g}$) and the sender provides information, then $\lambda_0^* = \frac{c}{v_g}$ due to strict concavity of the objective function. One can show that $(\lambda_0, \bar{\mu}_0, \underline{\mu}_0) = (\frac{c}{v_g}, 1, \frac{\mu_0 v_g - c}{v_g - c})$ is the only feasible information structure that satisfies (IR_0') and (F_0). Thus, the sender would provide information with $(\lambda_0, \bar{\mu}_0, \underline{\mu}_0) = (\frac{c}{v_g}, 1, \frac{\mu_0 v_g - c}{v_g - c})$ iff the sender surplus, $-K(\frac{c}{v_g}) + p \cdot \frac{c}{v_g}$, is positive (when it is 0, the sender is indifferent between providing information or not). Let $f(\tilde{c}) = -K(\frac{\tilde{c}}{v_g}) + p \cdot \frac{\tilde{c}}{v_g}$. We have $f(0) = 0$, f is strictly concave and obtains the maximum at $\tilde{c}^* = v_g\lambda_1^{**} < c < 1$. In addition, $f(v_g) < 0$ because $\lim_{\lambda \to 1} K'(\lambda) = +\infty$. Therefore, there exists a unique $\hat{c} \in (v_g\lambda_1^{**}, v_g)$ *s.t.* $f(c) \begin{cases} \geq 0, & if\ 0 \leq c \leq \hat{c} \\ < 0, & if\ c > \hat{c} \end{cases}$. Moreover, when the sender provides information, $\mu_0 \geq \frac{c}{v_g} \Rightarrow \hat{c} \leq \mu_0 v_g$. So, the sender does not provide information if $c > \hat{c}$ and provides information with $(\lambda_0^*, \bar{\mu}_0^*) = (\frac{c}{v_g}, 1)$ if $c < \hat{c}$. The receiver surplus is 0.

2. If $c \in [\mu_0 + v_b\lambda_1^{**}, v_g\lambda_1^{**})$ (i.e. $\lambda_1^{**} \geq \frac{\mu_0 - c}{-v_b} > \frac{c}{v_g}$) and the sender provides information, then $\lambda_0^* = \frac{\mu_0 - c}{-v_b}$ due to strict concavity of the objective function. One can show

31

that $(\lambda_0, \bar{\mu}_0, \underline{\mu}_0) = (\frac{\mu_0 - c}{-v_b}, \frac{-\mu_0 v_b}{\mu_0 - c}, 0)$ is the only feasible information structure that satisfies (IR_0') and (F_0). Thus, the sender would provide information with $(\lambda_0, \bar{\mu}_0, \underline{\mu}_0) = (\frac{\mu_0 - c}{-v_b}, \frac{-\mu_0 v_b}{\mu_0 - c}, 0)$ iff the sender surplus, $-K(\frac{\mu_0 - c}{-v_b}) + p \cdot \frac{\mu_0 - c}{-v_b}$, is positive. Since $-K(0) + p \cdot 0 = 0$, $\frac{\mu_0 - c}{-v_b} < \lambda_1^{**}$, and the objective function is strictly concave, the sender surplus is always strictly positive. So, the sender would always provide information. The receiver surplus is 0.

3. If $c < \mu_0 + v_b \lambda_1^{**} \wedge v_g \lambda_1^{**}$ (i.e. $\lambda_1^{**} \in \left(\frac{c}{v_g}, \frac{\mu_0 - c}{-v_b} \right)$), then the sender could obtain the maximum possible payoff by setting $(\lambda_0, \bar{\mu}_0) = (\lambda_1^{**}, \frac{\mu_0}{\lambda_1^{**}} \wedge 1)$. Let $\underline{\mu}_0 = \begin{cases} 0, \ if \ \mu_0 \le \lambda_1^{**} \\ \frac{\mu_0 - \lambda_1^{**}}{1 - \lambda_1^{**}}, \ if \ \mu_0 > \lambda_1^{**} \end{cases}$.

One can verify that $(\lambda_0, \bar{\mu}_0, \underline{\mu}_0)$ is feasible and satisfies (IR_0') and (F_0). We have shown in the proof of Lemma 1 that the sender surplus is strictly positive. So, the sender would provide information and $(\lambda_0^*, \bar{\mu}_0^*) = (\lambda_1^{**}, \frac{\mu_0}{\lambda_1^{**}} \wedge 1)$. The receiver surplus is

$$\begin{cases} \mu_0 + v_b \lambda_1^{**} - c, \ if \ \mu_0 \le \lambda_1^{**} \\ \lambda_1^{**} v_g - c, \ if \ \mu_0 > \lambda_1^{**} \end{cases} > 0.$$

If $c \ge v_g \lambda_1^{**}$, or $c < v_g \lambda_1^{**}$ and $\mu_1 \le c - v_b \lambda_1^{**}$, the expected receiver surplus in the second period is 0. The receiver incurs search cost without any immediate benefit in the first period under iterative signals. The expected receiver surplus in the first period is strictly negative if he searches. Therefore, he would not search, and iterative signals are not feasible. Now we consider the case in which $c < v_g \lambda_1^{**}$ and $\mu_1 > c - v_b \lambda_1^{**}$.

The sender's problem when she uses iterative signals is:

$$\Pi_{iter}(\mu_0) := \max -K(\lambda_0) + \lambda_0 \left[-K(\lambda_1^{**}) + p\lambda_1^{**} \right] \qquad (P_{iter})$$

$$\text{s.t. } \lambda_1(\bar{\mu}_1 + v_b) \ge \frac{1 + \lambda_0}{\lambda_0} c \qquad (IR_{0,iter})$$

$$\lambda_1(\bar{\mu}_1 + v_b) \ge c \qquad (IR_{1,iter})$$

$$\lambda_0 \bar{\mu}_0 + (1 - \lambda_0)\underline{\mu}_0 = \mu_0 \qquad (F_0)$$

$$\lambda_1 \bar{\mu}_1 + (1 - \lambda_1)\underline{\mu}_1 = \mu_1 \qquad (F_1)$$

$$\mu_1 = \bar{\mu}_0, \lambda_1 = \lambda_1^{**}$$

Note that $(IR_{0,iter})$ implies (IR_1^{iter}) and $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}})$ satisfies (F_1). $\forall \mu_1 \ge \lambda_1^{**}$, the optimal second-period strategy is always $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, 1)$. Therefore, choosing μ_1 above λ_1^{**} does not increase the second-period sender's payoff or relax the first-period constraints. So, we could restrict μ_1 to be less than or equal to λ_1^{**}.

i) $\mu_0 \ge c - v_b \lambda_1^{**}$

$$\Pi_1(\mu_0) = -K(\lambda_1^{**}) + p\lambda_1^{**} > \Pi_{iter}(\mu_0).$$

ii) $\mu_0 < c - v_b\lambda_1^{**}$

$$\Pi_1(\mu_0) = -K(\frac{\mu_0-c}{-v_b}) + \frac{(\mu_0-c)p}{-v_b}$$

$$(IR_{0,iter}) \Rightarrow \lambda_0 \geq \frac{c}{\mu_1 - c + v_b\lambda_1^{**}} (\Rightarrow \mu_1 \geq \frac{1+\lambda_0}{\lambda_0}c - v_b\lambda_1^{**}) \tag{1}$$

$$\geq \frac{c}{v_g\lambda_1^{**} - c}$$

$$(F_0) \Rightarrow \lambda_0 = \frac{\mu_0 - \underline{\mu}_0}{\mu_1 - \underline{\mu}_0} \leq \frac{\mu_0}{\mu_1} \overset{(1)}{\leq} \frac{\mu_0}{\frac{1+\lambda_0}{\lambda_0}c - v_b\lambda_1^{**}} \Rightarrow \lambda_0 \leq \frac{\mu_0 - c}{c - v_b\lambda_1^{**}}$$

A necessary condition for λ_0 to be well-defined is:

$$\frac{c}{v_g\lambda_1^{**} - c} \leq \frac{\mu_0 - c}{c - v_b\lambda_1^{**}} \Leftrightarrow \mu_0 \geq \frac{\lambda_1^{**}c}{v_g\lambda_1^{**} - c} (> \frac{c}{v_g})$$

Therefore, it is feasible for the sender to provide a one-period signal whenever it is feasible to provide iterative signals.

Define $\bar{\Pi}_{iter}(\mu_0) := \max_{0 \leq \lambda_0 \leq \frac{\mu_0-c}{c-v_b\lambda_1^{**}}} -K(\lambda_0) + \lambda_0 [-K(\lambda_1^{**}) + p\lambda_1^{**}]$. One can see that $\bar{\Pi}_{iter}(\mu_0) \geq \Pi_{iter}(\mu_0)$. Let $\lambda_0^*(\mu_0) := \underset{0 \leq \lambda_0 \leq \frac{\mu_0-c}{c-v_b\lambda_1^{**}}}{\operatorname{argmax}} -K(\lambda_0) + \lambda_0 [-K(\lambda_1^{**}) + p\lambda_1^{**}]$.

We want to show that:

$$\bar{\Pi}_{iter}(\mu_0) < \Pi_1(\mu_0)$$
$$\Leftrightarrow -K(\lambda_0^*(\mu_0)) + \lambda_0^*(\mu_0) [-K(\lambda_1^{**}) + p\lambda_1^{**}] < -K(\frac{\mu_0-c}{-v_b}) + \frac{\mu_0-c}{-v_b}p \tag{2}$$

Notice that

$$\frac{d}{d\lambda_0} \{-K(\lambda_0) + \lambda_0 [-K(\lambda_1^{**}) + p\lambda_1^{**}]\} = -K'(\lambda_0) - K(\lambda_1^{**}) + p\lambda_1^{**}$$

$$\Rightarrow \lambda_0^*(\mu_0) = \begin{cases} \frac{\mu_0-c}{c-v_b\lambda_1^{**}} & if \ \mu_0 \leq \mu_0^t \\ \frac{\mu_0^t-c}{c-v_b\lambda_1^{**}} & if \ \mu_0 > \mu_0^t \end{cases}$$

, where $\lambda_0^t > 0$ is defined by $-K'(\lambda_0^t) - K(\lambda_1^{**}) + p\lambda_1^{**} = 0$, $\mu_0^t = \lambda_0^t(c - v_b\lambda_1^{**}) + c$.

When $\mu_0 \leq \mu_0^t$, $\begin{cases} \lambda_0^*(\mu_0) = \frac{\mu_0-c}{c-v_b\lambda_1^{**}} > \frac{\mu_0-c}{-v_b} \Rightarrow -K(\lambda_0^*(\mu_0)) < -K(\frac{\mu_0-c}{-v_b}) \\ \lambda_0^*(\mu_0)p\lambda_1^{**} = \frac{\mu_0-c}{c-v_b\lambda_1^{**}}p\lambda_1^{**} < \frac{\mu_0-c}{-v_b}p & \Rightarrow \text{(2) holds,} \\ -\lambda_0^*(\mu_0)K(\lambda_1^{**}) < 0 \end{cases}$

where the first inequality holds because $-v_b > \mu_1 > c - v_b \lambda_1^{**}$.

When $\mu_0 > \mu_0^t$, $\bar{\Pi}_{iter}(\mu_0) = \bar{\Pi}_{iter}(\mu_0^t) < \Pi_1(\mu_0^t) < \Pi_1(\mu_0)$. So, (2) holds.

Thus, we have shown that $\Pi_1(\mu_0) > \bar{\Pi}_{iter}(\mu_0)$ for any μ_0 such that iterative signals are feasible. $\square$

Proof of Lemma 1. λ_0^{**} and λ_1^{**} are determined by the first order conditions: $-K'(\lambda_1^{**}) + p = 0$ and $-K'(\lambda_0^{**}) + p + K(\lambda_1^{**}) - p\lambda_1^{**} = 0$. $-K(0) + p \cdot 0 = 0$ & $-K'(\lambda) + p > 0$ for small $\lambda \Rightarrow -K(\lambda_1^{**}) + p\lambda_1^{**} > 0$. Therefore, $-K'(\lambda_0^{**}) + p + K(\lambda_1^{**}) - p\lambda_1^{**} = 0$ implies that $-K'(\lambda_0^{**}) + p > 0 = -K'(\lambda_1^{**}) + p, \Rightarrow K'(\lambda_0^{**}) < K'(\lambda_1^{**}) \Rightarrow \lambda_0^{**} < \lambda_1^{**}$. $-K(0) + p \cdot 0 + (1-0)[-K(\lambda_1^{**}) + p\lambda_1^{**}] > 0$ and strict concavity (w.r.t. λ_0) of the objective function imply that $-K(\lambda_0^{**}) + p\lambda_0^{**} + (1 - \lambda_0^{**})[-K(\lambda_1^{**}) + p\lambda_1^{**}] > 0$. $\square$

Proof of Proposition 2. It can be proved in the same way as in the proof of Proposition 1 when we derive the optimal one-period strategy of the sender, except that we need to replace the subscript 0 by 1 here. $\square$

Proof of Proposition 3.
(1) $v_g \lambda_1^{**} \le c < \widehat{c}$

If the sender provides information in both periods, the constrained program of the sender is:

$$\max -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right] \qquad (P_{2H})$$

$$\text{s.t. } (IR_0), (F_0), \mu_1 \ge \frac{c}{v_g}$$

We first transform (P_{2H}) into an equivalent constrained program that is easier to analyze.

Lemma 3. *Suppose $v_g \lambda_1^{**} \le c < \widehat{c}$. If $\mu_{0,1} \le \mu_0 < \frac{2v_g - c}{(v_g)^2}c$, the sender provides information in one period. If $\mu_0 \ge \frac{2v_g - c}{(v_g)^2}c$, (P_{2H}) is equivalent to:*

$$\max -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right] \qquad (P_{2H}')$$

$$\text{s.t. } \lambda_0 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - (1 + v_g)c}{-v_b v_g - c}\right]$$

Proof. The proof of the equivalence between (P_{2H}) and (P_{2H}') is similar to that of Lemma 2.

It is feasible for the sender to provide information at both periods if and only if the domain of λ_1 is non-empty: $\frac{c}{v_g} \le \frac{v_g\mu_0 - (2-p)c}{pv_g - c} \Leftrightarrow \mu_0 \ge \frac{2v_g - c}{(v_g)^2}c$. $\square$

Denote the optimal λ_0 without constraints by $\lambda_{0,H}^{**}$. $\lambda_{0,H}^{**} = \underset{\lambda_0}{\operatorname{argmax}} -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right]$. The following lemma summarizes the relative size of $\lambda_{0,H}^{**}$, λ_1^{**}, and $\frac{c}{v_g}$.

Lemma 4. $0 < \lambda_{0,H}^{**} < \lambda_1^{**} \le \frac{c}{v_g}$.

Proof. $\lambda_1^{**} < \frac{c}{v_g}$ is the assumption. F.O.C $\Rightarrow K'(\lambda_{0,H}^{**}) = p + K(\frac{c}{v_g}) - \frac{cp}{v_g}$. From Lemma 1, $K'(\lambda_1^{**}) = p$. $-K(\frac{c}{v_g}) + \frac{cp}{v_g} > 0$ when $c < \hat{c}$. So, $K'(\lambda_{0,H}^{**}) < K'(\lambda_1^{**}) \Rightarrow \lambda_{0,H}^{**} < \lambda_1^{**}$. $-K'(\lambda_{0,H}^{**}) + p + K(\frac{c}{v_g}) - \frac{cp}{v_g} = 0 \Rightarrow K'(\lambda_{0,H}^{**}) = p + K(\frac{c}{v_g}) - \frac{cp}{v_g} > p - \frac{cp}{v_g} = (1 - \frac{c}{v_g})p > 0$, where the last inequality follows from the assumption that $c < v_g$. Thus, $\lambda_{0,H}^{**} > 0$. $\lambda_1^{**} \le \frac{c}{v_g}$ is the assumption. $\qquad \square$

When it is feasible for the sender to provide information in both periods, $\mu_0 \ge \frac{2v_g - c}{(v_g)^2}c$, Lemma 4 and strict concavity of the objective function imply that the optimal two-period strategy of the sender is $(\lambda_t^*, \bar{\mu}_t^*) = (\frac{c}{v_g}, 1), t = 0, 1$. The sender surplus is $(2 - \frac{c}{v_g})\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right] > -K(\frac{c}{v_g}) + \frac{cp}{v_g}$, the sender surplus of the optimal one-period strategy. Therefore, the sender would always provide information in both periods as long as it is feasible.

(2) $c < v_g\lambda_1^{**}$

If the sender provides information in both periods, we first show that we could restrict the domain of μ_1 to be $\le \lambda_1^{**}$ given that $c < v_g\lambda_1^{**}$. The intuition is that $\forall \mu_1 \ge \lambda_1^{**}$, the optimal second-period strategy is always $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, 1)$. Therefore, choosing μ_1 above λ_1^{**} does not increase the second-period sender's payoff or relax the first-period constraints. So, we could restrict μ_1 to be less than or equal to λ_1^{**}. Formally, when $\lambda_1^{**} \le \mu_1 < \mu_0$, the constrained program of the sender is:

$$\max -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$$
$$\text{s.t. } \lambda_0(\bar{\mu}_0 + v_b) + (1 - \lambda_0)[\lambda_1^{**}v_g - c] \ge c \qquad (\tilde{IR}_0')$$
$$\lambda_0\bar{\mu}_0 + (1 - \lambda_0)\mu_1 = \mu_0 \qquad (F_0)$$
$$\mu_1 \in [\lambda_1^{**}, \mu_0)$$

$(\tilde{IR}_0')$ & $(F_0) \Rightarrow \lambda_0 \le \frac{-2c + \mu_0 - \mu_1 + v_g\lambda_1^{**}}{v_g\lambda_1^{**} - c - v_b - \mu_1} \le \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1 - \lambda_1^{**}) - c}$ ("$=$" when $\mu_1 = \lambda_1^{**}$). $(F_0) \Rightarrow \lambda_0 \ge \frac{\mu_0 - \mu_1}{1 - \mu_1}$. The domain of λ_0 is non-empty iff $\frac{\mu_0 - \mu_1}{1 - \mu_1} \le \frac{-2c + \mu_0 - \mu_1 + v_g\lambda_1^{**}}{v_g\lambda_1^{**} - c - v_b - \mu_1} \Leftrightarrow \mu_1 \le \frac{-2c + v_g\lambda_1^{**} + \mu_0[1 - v_g\lambda_1^{**} + c + v_b]}{v_g - c}$. Therefore, smaller μ_1 means it is more likely for the domain of λ_0 to be non-empty and larger upper bound of λ_0. So, the optimal μ_1 would never $\in (\lambda_1^{**}, \mu_0)$. Hence, the optimal strategy of the subgame is $(\lambda_1^*, \bar{\mu}_1^*) = \begin{cases} (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}}) & , \text{ if } \mu_1 \in (c - v_b\lambda_1^{**}, \lambda_1^{**}] \\ (\frac{\mu_1 - c}{-v_b}, \frac{-\mu_1 v_b}{\mu_1 - c}) & , \text{ if } \mu_1 \in [\frac{c}{v_g}, c - v_b\lambda_1^{**}] \end{cases}$ and the constrained

35

program of the full game is either[11]:

$$\max -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \qquad (P_{2S_+})$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \in [c - v_b\lambda_1^{**}, \lambda_1^{**}]$$

or:

$$\max -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] \qquad (P_{2S_0})$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \in [\frac{c}{v_g}, c - v_b\lambda_1^{**}]$$

1. S_+ strategy (solution to (P_{2S_+}))

   Proposition 11. *Suppose $c < v_g\lambda_1^{**}$ and $\mu_0 < \widehat{\mu}_0 = 2c - v_b\lambda_1^{**} - [c + (1 - \lambda_1^{**})v_b]\lambda_0^{**}$. If $\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, (P_{2S_+}) is feasible with the following solution. $\lambda_0^* = \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$; $\bar{\mu}_0^* = \begin{cases} \frac{(v_b-c)(c-v_b\lambda_1^{**})-v_b\mu_0}{\mu_0-2c+v_b\lambda_1^{**}} \in (-v_b, 1) & ,if\ \widehat{\mu}_1(\mu_0) < c - v_b\lambda_1^{**} \\ 1 & ,if\ \widehat{\mu}_1(\mu_0) \geq c - v_b\lambda_1^{**} \end{cases}$; $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}})$; $\mu_1^* = \widehat{\mu}_1(\mu_0) \vee c - v_b\lambda_1^{**}$, where $\widehat{\mu}_1(\mu_0) = \frac{2c-v_b\lambda_1^{**}-(1+c-v_b\lambda_1^{**}+v_b)\mu_0}{c-v_b-\mu_0}$. The receiver gets zero surplus.*

   Proof. We first transform (P_{2S_+}) into an equivalent constrained program that is easier to analyze.

   Lemma 5. *Suppose $c < v_g\lambda_1^{**}$ and $\mu_0 \geq \mu_{0,1}$. If $\mu_0 \leq 2c - v_b\lambda_1^{**}$ or $\mu_0 < \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, the sender provides information in one period. If $\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, (P_{2S_+}) is equivalent to:*

$$\max -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \qquad (P_{2S_+}'')$$
$$\text{s.t. } \lambda_0 \in \left(0, \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}\right]$$

   Proof. We first show that, if $\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, (P_{2S_+}) is equivalent

---

[11] We include $\mu_1 = c - v_b\lambda_1^{**}$ in (P_{2S_+}) as well to simplify the exposition.

to:

$$\max -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \qquad (P'_{2S_+})$$

$$\text{s.t. } \lambda_0 \in \left[\frac{\mu_0 - \mu_1}{1 - \mu_1}, \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1 - \lambda_1^{**}) - c} \wedge \frac{\mu_0 - \mu_1}{-v_b - \mu_1}\right]$$

$$\mu_1 \in [c - v_b\lambda_1^{**} \vee \widehat{\mu_1}(\mu_0), \lambda_1^{**}]$$

$$, \text{ where } \widehat{\mu_1}(\mu_0) = \frac{2c - v_b\lambda_1^{**} - (1 + c - v_b\lambda_1^{**} + v_b)\mu_0}{c - v_b - \mu_0}$$

$(F_0) \Rightarrow \lambda_0 = \frac{\mu_0 - \mu_1}{\bar{\mu}_0 - \mu_1} \in \left[\frac{\mu_0 - \mu_1}{1 - \mu_1}, \frac{\mu_0 - \mu_1}{-v_b - \mu_1}\right]$. (IR_0) & $(F_0) \Rightarrow \lambda_0 \le \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1 - \lambda_1^{**}) - c}$. For λ_0 to be positive, we need $\mu_0 > 2c - v_b\lambda_1^{**}$. The domain of λ_0 is non-empty iff $\frac{\mu_0 - \mu_1}{1 - \mu_1} \le \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1 - \lambda_1^{**}) - c} \Leftrightarrow \mu_1 \ge \widehat{\mu_1}(\mu_0)$. For $\mu_1 \le \lambda_1^{**}$, we need $\widehat{\mu_1}(\mu_0) \le \lambda_1^{**} \Leftrightarrow \mu_0 \ge \frac{(2 - \lambda_1^{**})c}{v_g(1 - \lambda_1^{**}) + c}$. We also have that $\mu_1 = \underline{\mu}_0 < \mu_0$. Thus, the constraints in (P_{2S_+}) imply the constraints in (P'_{2S_+}).

For any (λ_0, μ_1) satisfying the constraints in (P'_{2S_+}), consider $(\lambda_0, \mu_1, \bar{\mu}_0 = \frac{\mu_0 - (1 - \lambda_0)\mu_1}{\lambda_0}, \bar{\mu}_1 = \frac{\mu_1}{\lambda_1^{**}} \wedge 1, \underline{\mu}_1 = \frac{\mu_1 - \lambda_1^{**}\bar{\mu}_1}{1 - \lambda_1^{**}})$. (IR_0) & (F_0) are satisfied by construction. $\bar{\mu}_0 = \frac{\mu_0 - (1 - \lambda_0)\mu_1}{\lambda_0} > \frac{\mu_0 - (1 - \lambda_0)\mu_0}{\lambda_0} = \mu_0$. $\bar{\mu}_0 = \frac{\mu_0 - (1 - \lambda_0)\mu_1}{\lambda_0} \le \frac{\mu_0 - (1 - \frac{\mu_0 - \mu_1}{1 - \mu_1})\mu_1}{\frac{\mu_0 - \mu_1}{1 - \mu_1}} = 1$. One can verify that $\bar{\mu}_1 \in (-v_b, 1]$, $\underline{\mu}_1 \in [0, -v_b)$. Therefore, the $(\lambda_0, \mu_1, \bar{\mu}_0, \bar{\mu}_1, \underline{\mu}_1)$ we constructed satisfies all the constraints in (P_{2S_+}) and is feasible.

Therefore, the two programs are equivalent.

We then show that (P'_{2S_+}) is equivalent to (P''_{2S_+}). It is clear that the constraints in (P'_{2S_+}) imply the constraints in (P''_{2S_+}). We now show that for any $\lambda_0 \in \left(0, \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1 - \lambda_1^{**}) - c}\right]$, we could find (λ_0, μ_1) that satisfies the constraints in (P'_{2S_+}) and is feasible. Since $\lambda_0 = \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1 - \lambda_1^{**}) - c}$ maximizes the objective function among $\lambda_0 \in \left(0, \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1 - \lambda_1^{**}) - c}\right]$ when $\mu_0 < \widehat{\mu}_0$, we only need to verify (by construction) that $\lambda_0 = \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1 - \lambda_1^{**}) - c}$ can be obtained.

i) $\widehat{\mu_1}(\mu_0) < c - v_b\lambda_1^{**}$

Consider $\mu_1 = c - v_b\lambda_1^{**}, \lambda_0 = \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1 - \lambda_1^{**}) - c}, \bar{\mu}_0 = \frac{\mu_0 - (1 - \lambda_0)\mu_1}{\lambda_0} = \frac{(v_b - c)(c - v_b\lambda_1^{**}) - v_b\mu_0}{\mu_0 - 2c + v_b\lambda_1^{**}}$. By construction, (IR_0) and (F_0) are satisfied; μ_1's constraints are also satisfied. So, we just need to verify that $\bar{\mu}_0 \in (p, 1)$. $\bar{\mu}_0 < 1 \Leftrightarrow (v_b - c)(c - v_b\lambda_1^{**}) - v_b\mu_0 < \mu_0 - 2c + v_b\lambda_1^{**} \Leftrightarrow \widehat{\mu_1}(\mu_0) < c - v_b\lambda_1^{**}$, which is the assumption. $\bar{\mu}_0 > -v_b \Leftrightarrow c < -v_b(1 - \lambda_1^{**})$, which holds because $\mu_0 > 2c - v_b\lambda_1^{**} \Rightarrow c < \frac{1}{2}(\mu_0 + v_b\lambda_1^{**}) \le \mu_0 + v_b\lambda_1^{**} < -v_b + v_b\lambda_1^{**} = -v_b(1 - \lambda_1^{**})$.

ii) $\widehat{\mu_1}(\mu_0) \ge c - v_b\lambda_1^{**}$

Consider $\lambda_0 = \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, \bar{\mu}_0 = 1, \mu_1 = \frac{\mu_0 - \lambda_0\bar{\mu}_0}{1-\lambda_0} = \widehat{\mu}_1(\mu_0)$. By construction, (IR_0) and (F_0) are satisfied; μ_1's constraints are also satisfied. So, we just need to verify that $\mu_1 = \widehat{\mu}_1(\mu_0) \in [c - v_b\lambda_1^{**}, \lambda_1^{**}]$. $\widehat{\mu}_1(\mu_0) \geq c - v_b\lambda_1^{**}$ is the assumption. $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c} \Rightarrow \widehat{\mu}_1(\mu_0) \leq \lambda_1^{**}$.

$\square$

When $\frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c} \geq \lambda_0^{**} (\Leftrightarrow \mu_0 \geq \widehat{\mu}_0)$, the optimal λ_0 is λ_0^{**}. When $\frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c} < \lambda_0^{**}$, the optimal λ_0 is $\frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$ due to strict concavity of the objective function (denote it by $J(\lambda_0)$). Since the one-period optimal sender surplus is $-K(\lambda_1^{**}) + p\lambda_1^{**} = J(0)$, $J(\cdot)$ is strictly concave and obtains the unique maximum value at $\lambda_0^{**} > \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$, we have $J(\frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}) > J(0)$. So, the sender always provides information in both periods when it is feasible ($\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$). We will use this observation in the later proofs. According to the proof of Lemma 5, the receiver always gets zero surpluses when $\mu_0 < \widehat{\mu}_0$. $\mu_1 = \widehat{\mu}_1(\mu_0) \vee c - v_b\lambda_1^{**}$ is the smallest μ_1 that supports $\lambda_0^* = \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$, which gives the receiver the largest surplus in the first period. So, $\mu_1^* = \widehat{\mu}_1(\mu_0) \vee c - v_b\lambda_1^{**}$. $(F_0) \Rightarrow \bar{\mu}_0^* = \frac{\mu_0 - (1-\lambda_0)\mu_1}{\lambda_0} =$

$$\begin{cases} \frac{(v_b-c)(c-v_b\lambda_1^{**})-v_b\mu_0}{\mu_0 - 2c + v_b\lambda_1^{**}} \in (p,1) & , if \ \widehat{\mu}_1(\mu_0) < c - v_b\lambda_1^{**} \\ 1 & , if \ \widehat{\mu}_1(\mu_0) \geq c - v_b\lambda_1^{**} \end{cases}$$

$\square$

2. S_0 strategy (solution to (P_{2S_0}))

Proposition 12. *Suppose $c \leq v_g\lambda_0^{**}$. When $\mu_0 \geq \frac{2v_g-c}{(v_g)^2}c$, (P'_{2S_0}) is feasible. λ_0^*, λ_1^*, and μ_1^* are continuous and increase in μ_0, while $\bar{\mu}_0^*$ and $\bar{\mu}_1^*$ are continuous and decrease in μ_0, in the solution to (P'_{2S_0}). The receiver gets zero surplus at each period.*

Proof. We first transform (P_{2S_0}) into an equivalent constrained program that is easier to analyze.

Lemma 6. *Suppose $c < v_g\lambda_1^{**}$. If $\mu_{0,1} \leq \mu_0 < \frac{2v_g-c}{(v_g)^2}c$, the sender provides information in one period. If $\mu_0 \geq \frac{2v_g-c}{(v_g)^2}c$, (P_{2S_0}) is equivalent to:*

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] \qquad (P'_{2S_0})$$

$$s.t. \ \lambda_0 \in \left[\frac{\mu_0 - \mu_1}{1-\mu_1}, \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}\right]$$

$$\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right]$$

38

Proof. Using the same argument as the proof of Lemma 5, one can show that (P_{2S_0}) is equivalent to:

$$\max -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] \qquad (P''_{2S_0})$$

$$\text{s.t. } \lambda_0 \in \left[\frac{\mu_0 - \mu_1}{1 - \mu_1}, \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}\right]$$

$$\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c} \wedge c - v_b\lambda_1^{**}\right]$$

We just need to show that (P''_{2S_0}) is equivalent to (P'_{2S_0}).

If $\frac{v_g\mu_0 - c}{v_g - c} \leq c - v_b\lambda_1^{**}$, μ_1's constraint becomes $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right]$. So, (P''_{2S_0}) is equivalent to (P'_{2S_0}).

If $\frac{v_g\mu_0 - c}{v_g - c} > c - v_b\lambda_1^{**}$, denote the solution to (P'_{2S_0}) by (λ_0, μ_1).

(a) $(\lambda_0^{**}, \lambda_1^{**})$ could be obtained $(\lambda_1 = \lambda_1^{**} \Leftrightarrow \mu_1 = c - v_b\lambda_1^{**})$ in (P''_{2S_0})

$\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c} \wedge c - v_b\lambda_1^{**}\right]$ is equivalent to $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right]$, as the optimal μ_1 under the latter (relaxed) constraint would be $c - v_b\lambda_1^{**}$. So, (P''_{2S_0}) is equivalent to (P'_{2S_0}).

(b) $(\lambda_0^{**}, \lambda_1^{**})$ could not be obtained in (P''_{2S_0})

Suppose $\mu_1 > c - v_b\lambda_1^{**}$. If $\lambda_0 > \frac{\mu_0 - \mu_1}{1 - \mu_1}$, consider $(\lambda_0' = \lambda_0, \mu_1' = \mu_1 - \varepsilon)$. For small enough ε, it is feasible and gives the sender a strictly higher payoff. A contradiction! If $\lambda_0 = \frac{\mu_0 - \mu_1}{1 - \mu_1}$ instead, we have $\frac{\mu_0 - \mu_1}{1 - \mu_1} \geq \lambda_0^{**}$. A contradiction!

Therefore, $\mu_1 \leq c - v_b\lambda_1^{**}$ and thus (P''_{2S_0}) is equivalent to (P'_{2S_0}).

In sum, (P''_{2S_0}) is equivalent to (P'_{2S_0}). □

To solve (P'_{2S_0}), we introduce another lemma:

Lemma 7. *Suppose $c < v_g\lambda_1^{**}$. For $\mu_0 < \widehat{\mu}_0$, λ_0 is binding at the upper bound in the solution to (P'_{2S_0}).*

Proof. To solve (P'_{2S_0}), we consider several cases.

i) $\lambda_0 \leq \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}$ is binding and μ_1's constraints are not binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] + \eta\left(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1} - \lambda_0\right)$

s.t. $\eta \geq 0, \eta\left(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1} - \lambda_0\right) = 0$.

$$\text{F.O.C.} \Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{v_b} - \eta = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1-c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] + \eta \cdot \frac{\mu_0+v_b-c}{(v_b+\mu_1)^2} = 0 \end{cases}$$

Plug in $\lambda_0 = \frac{\mu_0-\mu_1-c}{-v_b-\mu_1}$. Dividing the second equality by $\frac{\mu_0+v_b-c}{(v_b+\mu_1)^2}$ and comparing with the first equality, we obtain:

$$\eta = -(v_b + \mu_1)\left[K'(\frac{\mu_1-c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right]$$

$$= -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{v_b}$$

$$\Rightarrow K(\frac{\mu_1-c}{-v_b}) + \frac{v_b+\mu_1}{v_b}K'(\frac{\mu_1-c}{-v_b}) - K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) - \frac{cp}{v_b} = 0 \qquad (*)$$

$\frac{\partial}{\partial \mu_1}\left[K(\frac{\mu_1-c}{-v_b}) + \frac{v_b+\mu_1}{v_b}K'(\frac{\mu_1-c}{-v_b})\right] = -\frac{v_b+\mu_1}{(v_b)^2}K''(\frac{\mu_1-c}{-v_b}) > 0$. So, the sum of the first two terms of the LHS of $(*)$ strictly increases in μ_1. $\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}$ strictly decreases in μ_1, $K'(\cdot)$ strictly increases in μ_1. So, $-K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})$ strictly increases in μ_1. Thus, the LHS of $(*)$ strictly increases in μ_1. When μ_0 increases, the LHS of $(*)$ is strictly negative if μ_1 is unchanged. Therefore, μ_1 also has to increase. So, the sum of the first two terms of the LHS of $(*)$ increases. As a result, the third term, $-K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) = -K'(\lambda_0)$ has to decrease strictly. So, λ_0 has to increase strictly. In sum, the optimal λ_0 and μ_1 are strictly increasing in μ_0.

ii) $\mu_1 \leq \frac{v_g\mu_0-c}{v_g-c}$ is binding.

When $\mu_1 = \frac{v_g\mu_0-c}{v_g-c}$, $\lambda_0 \in \{\frac{c}{v_g}\}$. So, λ_0 is binding at the upper bound.

iii) $\mu_1 \geq \frac{c}{v_g}$ is binding and λ_0 is not binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] + \eta\left(\mu_1 - \frac{c}{v_g}\right)$

$s.t. \ \eta \geq 0, \eta\left(\mu_1 - \frac{c}{v_g}\right) = 0$.

$$\text{F.O.C.} \Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{v_b} = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1-c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] + \eta = 0 \end{cases}$$

The second equality $\Rightarrow \eta = -\frac{1-\lambda_0}{v_b}\left[K'(\frac{\mu_1-c}{-v_b}) - p\right] \overset{c < v_g\lambda_1^{**}}{<} -\frac{1-\lambda_0}{v_b}\left[K'(\lambda_1^{**}) - p\right] = 0$. But $\eta \geq 0$. A contradiction! So, this case cannot happen.

iv) $\mu_1 \geq \frac{c}{v_g}$ is binding and $\lambda_0 \geq \frac{\mu_0-\mu_1}{1-\mu_1}$ is binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] + \eta\left(\mu_1 - \frac{c}{v_g}\right) + \xi\left(\lambda_0 - \frac{\mu_0-\mu_1}{1-\mu_1}\right) \ s.t. \ \eta \geq 0, \eta\left(\mu_1 - \frac{c}{v_g}\right) = 0, \xi \geq 0, \xi\left(\lambda_0 - \frac{\mu_0-\mu_1}{1-\mu_1}\right) = 0$.

40

$$\text{F.O.C.} \Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{v_b} + \xi = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1-c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] + \eta + \xi\frac{1-\mu_0}{(1-\mu_1)^2} = 0 \end{cases}$$

Similar to the previous case, the LHS of the second equality > 0. A contradiction! So, this case cannot happen.

v) $\lambda_0 \geq \frac{\mu_0-\mu_1}{1-\mu_1}$ is binding and μ_1 is not binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] + \xi\left(\lambda_0 - \frac{\mu_0-\mu_1}{1-\mu_1}\right)$

s.t. $\xi \geq 0, \xi\left(\lambda_0 - \frac{\mu_0-\mu_1}{1-\mu_1}\right) = 0$.

$$\text{F.O.C.} \Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b} + \xi = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1-c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] + \xi\frac{1-\mu_0}{(1-\mu_1)^2} = 0 \end{cases}$$

Similar to the previous case, the LHS of the second equality > 0. A contradiction! So, this case cannot happen.

vi) both λ_0 and μ_1 are not binding.

The solution is the unconstrained optimal solution $(\lambda_0^{**}, \lambda_1^{**})$. But we have assumed that it is not feasible.

i) to vi) finish the proof of Lemma 7. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

According to Lemma 7, if $\mu_0 \geq \frac{2v_g-c}{(v_g)^2}c$, (P'_{2S_0}) is equivalent to:

$$\max -K(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + p\frac{\mu_0-\mu_1-c}{-v_b-\mu_1} + (1 - \frac{\mu_0-\mu_1-c}{-v_b-\mu_1})\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right]$$
$$(P'''_{2S_0})$$

s.t. $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c}\right]$

The first order derivative of the objective function w.r.t. μ_1 is:

$$D(\mu_0, \mu_1)$$
$$:= \frac{\partial}{\partial\mu_1}\left\{-K(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + p\frac{\mu_0-\mu_1-c}{-v_b-\mu_1} + (1 - \frac{\mu_0-\mu_1-c}{-v_b-\mu_1})\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right]\right\}$$
$$= \frac{\mu_0+v_b-c}{(\mu_1+v_b)^2}\left[K(\frac{\mu_1-c}{-v_b}) + \frac{v_b+\mu_1}{v_b}K'(\frac{\mu_1-c}{-v_b}) - K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) - \frac{cp}{v_b}\right]$$

The first term of $D(\mu_0, \mu_1)$, $\frac{\mu_0+v_b-c}{(\mu_1+v_b)^2}$, is always strictly negative. The second term, $K(\frac{\mu_1-c}{-v_b}) + \frac{v_b+\mu_1}{v_b}K'(\frac{\mu_1-c}{-v_b}) - K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) - \frac{cp}{v_b}$, is the LHS of $(*)$, which has been shown

to be strictly increasing in μ_1 in the proof of Lemma 7. One can see that $D(\mu_0, \mu_1)$ is strictly negative when μ_1 is large. Thus, $D(\mu_0, \mu_1)$ is always negative or positive for μ_1 small and negative for μ_1 large. Let $\mu_1^{**}(\mu_0)$ be the cutoff value such that $D(\mu_0, \mu_1) \geq 0$ for $\mu_1 \leq \mu_1^{**}(\mu_0)$ and $D(\mu_0, \mu_1) \leq 0$ for $\mu_1 \geq \mu_1^{**}(\mu_0)$ ($\mu_1^{**}(\mu_0) := -\infty$ if $D(\mu_0, \mu_1)$ is always negative). Since $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g \mu_0 - c}{v_g - c}\right]$, the optimal $\mu_1^*(\mu_0) = \left[\frac{c}{v_g} \vee \mu_1^{**}(\mu_0)\right] \wedge \frac{v_g \mu_0 - c}{v_g - c}$.

One can see that we could define $\tilde{\mu}_1^{**}(\mu_0) := \begin{cases} \frac{c}{v_g}, & if \ D(\mu_0, \mu_1) \text{ is always negative} \\ \mu_1^{**}(\mu_0), & \text{otherwise} \end{cases}$.

$\tilde{\mu}_1^{**}(\mu_0) \in (-\infty, +\infty)$ and $\mu_1^*(\mu_0) = \left[\frac{c}{v_g} \vee \tilde{\mu}_1^{**}(\mu_0)\right] \wedge \frac{v_g \mu_0 - c}{v_g - c}$. Since $\tilde{\mu}_1^{**}(\mu_0)$ is continuous in μ_0, $\mu_1^*(\mu_0)$ is also continuous in μ_0. It then implies that $\lambda_0^*(\mu_0) = \frac{\mu_0 - \mu_1^*(\mu_0) - c}{-v_b - \mu_1^*(\mu_0)}$, $\lambda_1^*(\mu_0) = \frac{\mu_1^*(\mu_0) - c}{-v_b}$, and $\bar{\mu}_1^*(\mu_0) = \frac{-\mu_1^*(\mu_0)v_b}{\mu_1^*(\mu_0) - c}$ are continuous in μ_0.

We have shown in the proof of Lemma 7 that λ_0^* and μ_1^* strictly increase in μ_0 when μ_1^* is the interior solution. Now we consider the case when μ_1 is binding. When $\mu_1^* = \frac{c}{v_g}$, $\lambda_0^* = \frac{\mu_0 - \mu_1^* - c}{-v_b - \mu_1^*} = \frac{\mu_0 - \frac{c}{v_g} - c}{-v_b - \frac{c}{v_g}}$ strictly increases in μ_0. When $\mu_1^* = \frac{v_g \mu_0 - c}{v_g - c}$, it is strictly increasing in μ_0 and $\lambda_0^* = \frac{\mu_0 - \mu_1^* - c}{-v_b - \mu_1^*} = \frac{c}{v_g}$ is constant. Together with the continuity property we just established, we have shown that λ_0^* and μ_1^* are (weakly) increasing in μ_0. Thus, $\lambda_1^* = \frac{\mu_1^* - c}{-v_b}$ is (weakly) increasing in μ_0 and $\bar{\mu}_1^* = \frac{-\mu_1^* v_b}{\mu_1^* - c}$ is (weakly) decreasing in μ_0. $\qquad\square$

According to the proof of Proposition 1 and Lemma 6, the sender does not provide information iff $\mu_0 < \mu_{0,1}$ and provide information in one period if $\mu_{0,1} \leq \mu_0 < \frac{2v_g - c}{(v_g)^2}c$. Thus, we just need to determine whether she provides information in one period or in both periods when $\mu_0 \geq \frac{2v_g - c}{(v_g)^2}c$ by comparing the sender surplus of the optimal one-period strategy and the optimal sender surplus of the S_0 strategy.

(a) $c \leq v_g \lambda_0^{**}$

Define $\mu_{1,2} := \inf\{\mu_0 \geq \frac{2v_g - c}{(v_g)^2}c : \Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0)\}$. One can see that $\mu_{1,2} \in [\frac{2v_g - c}{(v_g)^2}c, \widehat{\mu_0})$ and $\Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2})$. According to Lemma 7,

$$\Pi_1(\mu_0) = -K(\frac{\mu_0 - c}{-v_b} \wedge \lambda_1^{**}) + p(\frac{\mu_0 - c}{-v_b} \wedge \lambda_1^{**})$$

$$\Pi_{S_0}(\mu_0)$$
$$= \max_{\mu_1} -K(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}) + p\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1} + (1 - \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1})\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right]$$
$$s.t \ \mu_1 \in \left[\frac{c}{v_g}, \frac{v_g \mu_0 - c}{v_g - c}\right]$$

42

i) $\mu_{1,2} \geq c - v_b\lambda_1^{**}$

$\forall \mu_0 \in (\mu_{1,2}, \widehat{\mu}_0], \Pi_{S_0}(\mu_0) > \Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2}) = \Pi_1(\mu_0).$

ii) $\mu_{1,2} < c - v_b\lambda_1^{**}$

$\forall \mu_0 \in [\mu_{1,2}, c - v_b\lambda_1^{**}), \frac{d\Pi_1(\mu_0)}{d\mu_0} = K'(\frac{\mu_0-c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b}.$

A. $\mu_1(\mu_0) = \mu_1^u(\mu_0) = \frac{v_g\mu_0-c}{v_g-c}$

$\Pi_{S_0}(\mu_0) = -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})[-K(\frac{\mu_1^u(\mu_0)-c}{-v_b}) + \frac{(\mu_1^u(\mu_0)-c)p}{-v_b}].$

For $\Delta > 0$ small enough, we have $\mu_0 + \delta < c - v_b\lambda_1^{**}$, $\forall \delta \in (0, \Delta)$. Consider $\mu_{0,\delta} = \mu_0 + \delta \in (\mu_0, \mu_0 + \Delta)$, we have $\Pi_{S_0}(\mu_{0,\delta}) \geq \underline{\Pi}_{S_0}(\mu_{0,\delta}) :=$ $-K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})[-K(\frac{\mu_1^u(\mu_{0,\delta})-c}{-v_b}) + \frac{(\mu_1^u(\mu_{0,\delta})-c)p}{-v_b}].$ Noticing that $\Pi_{S_0}(\mu_0) = \underline{\Pi}_{S_0}(\mu_0)$, we have

$$\frac{d\Pi_{S_0}(\mu_0)}{d\mu_0} \geq \frac{d\underline{\Pi}_{S_0}(\mu_0)}{d\mu_0} = K'(\frac{\mu_1^u(\mu_0)-c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b} > \frac{d\Pi_1(\mu_0)}{d\mu_0}$$

So, $\Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0), \forall \mu_0 \in [\mu_{1,2}, c - v_b\lambda_1^{**})$, and the inequality is strict when $\mu_0 > \mu_{1,2}$.

B. $\mu_1(\mu_0) < \mu_1^u(\mu_0)$

Let $\lambda_0 = \frac{\mu_0-\mu_1(\mu_0)-c}{-v_b-\mu_1(\mu_0)}$. For $\Delta > 0$ small enough, we have $\mu_0 + \delta < c - v_b\lambda_1^{**}$ and $\mu_1(\mu_0) + \frac{\delta}{1-\lambda_0} < \mu_1^u(\mu_0) < \mu_1^u(\mu_0 + \delta)$, $\forall \delta \in (0, \Delta)$. Consider $\mu_{0,\delta} = \mu_0 + \delta \in (\mu_0, \mu_0 + \Delta)$, noticing that $\frac{\mu_0+\delta-(\mu_1+\frac{\delta}{1-\lambda_0})-c}{-v_b-(\mu_1+\frac{\delta}{1-\lambda_0})} = \lambda_0$, we have

$\Pi_{S_0}(\mu_{0,\delta}) \geq \widetilde{\Pi}_{S_0}(\mu_{0,\delta}) := -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)[-K(\frac{\mu_1(\mu_0)+\frac{\delta}{1-\lambda_0}-c}{p}) + \frac{(\mu_1(\mu_0)+\frac{\delta}{1-\lambda_0}-c)p}{-v_b}].$

$$\frac{d\Pi_{S_0}(\mu_0)}{d\mu_0} \geq \frac{d\widetilde{\Pi}_{S_0}(\mu_0)}{d\mu_0}$$

$$= (1 - \lambda_0)\left[-\frac{1}{p}K'(\frac{\mu_1(\mu_0)-c}{p})\frac{1}{1-\lambda_0} + \frac{1}{1-\lambda_0}\right]$$

$$= 1 - \frac{1}{p}K'(\frac{\mu_1(\mu_0)-c}{p})$$

$$\geq 1 - \frac{1}{p}K'(\frac{\mu_1^u(\mu_0)-c}{p})$$

$$> \frac{d\Pi_1(\mu_0)}{d\mu_0}$$

So, $\Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0)$ and the inequality is strict when $\mu_0 > \mu_{1,2}$.

In sum, $\Pi_{S_0}(\mu_0) > \Pi_1(\mu_0)$, $\forall \mu_0 \in (\mu_{1,2}, \widehat{\mu}_0].$

(b) $v_g\lambda_0^{**} < c < v_g\lambda_1^{**}$

The following lemma provides a closed-form solution to program (P_{2S_0}) when the search cost is intermediate:

Lemma 8. *Suppose $v_g\lambda_0^{**} < c < v_g\lambda_1^{**}$. $\lambda_0^* = \frac{c}{v_g}, \mu_1^* = \frac{v_g\mu_0-c}{v_g-c}$ in the solution to (P_{2S_0}).*

Proof. According to the proof of Proposition 12, λ_0^* is binding at the upper bound and increases in μ_0 in the solution to (P'_{2S_0}), for $c < v_g\lambda_1^{**}$. One can see that $\lambda_0^* = \frac{c}{v_g}$ and $\mu_1^* = c - v_b\lambda_1^{**}$ for μ_0 large enough in the solution to (P'_{2S_0}). Because $\lambda_0^* \geq \frac{c}{v_g}$, the only way for λ_0^* to be increasing in μ_0 is for it to always be $\frac{c}{v_g}$. Given $\lambda_0^* = \frac{c}{v_g}$, the optimal $\mu_1^* = \frac{v_g\mu_0-c}{v_g-c}$ for (P'_{2S_0}). Lemma 6 shows that (P'_{2S_0}) is equivalent to (P_{2S_0}). So, $\lambda_0^* = \frac{c}{v_g}, \mu_1^* = \frac{v_g\mu_0-c}{v_g-c}$ are also the solutions to (P_{2S_0}). $\square$

Define $\mu_{1,2} := \inf\{\mu_0 \geq \frac{2v_g-c}{(1-p)^2}c : \Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0)$ or S_+ strategy is feasible$\}$. Note that $\mu_{1,2} \leq \mu_{2,+}$. If S_+ strategy is feasible $\forall \mu_0 > \mu_{1,2}$, the 1-period sender surplus would always be dominated by the 2-period sender surplus $\forall \mu_0 > \mu_{1,2}$, as the optimal S_+ strategy generates a strictly higher sender surplus than the optimal 1-period strategy. We now consider the case in which S_+ strategy is not feasible for some $\mu_0 > \mu_{1,2}$, which implies that $\Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2})$.

$$\Pi_1(\mu_0) = - K(\frac{\mu_0 - c}{-v_b} \wedge \lambda_1^{**}) + p(\frac{\mu_0 - c}{-v_b} \wedge \lambda_1^{**})$$

$$\Pi_{S_0}(\mu_0) = - K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})\left[-K(\frac{\mu_1^u(\mu_0) - c}{-v_b}) + \frac{(\mu_1^u(\mu_0) - c)p}{-v_b}\right]$$

i) $\mu_{1,2} \geq c - v_b\lambda_1^{**}$

   $\forall \mu_0 \in (\mu_{1,2}, \mu_{2,+}], \Pi_{S_0}(\mu_0) > \Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2}) = \Pi_1(\mu_0).$

ii) $\mu_{1,2} < c - v_b\lambda_1^{**}$

   $\forall \mu_0 \in [\mu_{1,2}, c - v_b\lambda_1^{**}],$

$$\frac{d\Pi_1(\mu_0)}{d\mu_0} = K'(\frac{\mu_0 - c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b}$$

$$\frac{d\Pi_{S_0}(\mu_0)}{d\mu_0} = K'(\frac{\mu_1^u(\mu_0) - c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b} > \frac{d\Pi_1(\mu_0)}{d\mu_0}$$

   So, $\Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0)$ and the inequality is strict when $\mu_0 > \mu_{1,2}$. $\Pi_{S_0}(c - v_b\lambda_1^{**}) > \Pi_1(c - v_b\lambda_1^{**})$. $\forall \mu_0 \in [c - v_b\lambda_1^{**}, \mu_{2,+}], \Pi_{S_0}(\mu_0) \geq \Pi_{S_0}(c - v_b\lambda_1^{**}) > \Pi_1(c - v_b\lambda_1^{**}) = \Pi_1(\mu_0).$

In sum, $\forall \mu_0 \in (\mu_{1,2}, \mu_{2,+}], \Pi_{S_0}(\mu_0) > \Pi_1(\mu_0)$.

One can see that the optimal S_+ strategy always generates a strictly higher (and strictly positive) sender surplus than the optimal 1-period strategy. Therefore, the sender always provides information in both periods when S_+ strategy is feasible.[12] By Lemma 5, S_+ strategy is feasible iff $\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$ when $c < v_g\lambda_1^{**}$. Hence, together with the above results on S_0 strategy, there exists $\mu_{1,2} \in [\frac{2v_g-c}{(v_g)^2}c, 2c - v_b\lambda_1^{**} \vee \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}]$ such that the sender does not provide information if $\mu_0 < \mu_{0,1}$, provides information in one period if $\mu_0 \in [\mu_{0,1}, \mu_{1,2})$, and provides information in both periods if $\mu_0 > \mu_{1,2}$. $\square$

Proof of Proposition 4.

(1) High Search Cost $(v_g\lambda_1^{**} \leq c < \widehat{c})$

   It has been shown in the proof of Proposition 3. Since the optimal strategy does not depend on the prior, the sender's payoff does not depend on the prior either.

(2) Low Search Cost $(c \leq \tilde{c} = v_g K'^{-1}\left[\frac{K(\lambda_1^{**})}{\lambda_1^{**}}\right])$

   We first prove the following claim:

   Lemma 9. $\tilde{c} < v_g\lambda_0^{**} < v_g\lambda_1^{**}$.

   Proof. We have $K'(\frac{\tilde{c}}{v_g}) = \frac{K(\lambda_1^{**})}{\lambda_1^{**}}$.

   $$\text{FOC of } (P_b) \Rightarrow K'(\lambda_0^{**}) = K(\lambda_1^{**}) + p(1 - \lambda_1^{**})$$
   $$\Rightarrow \lambda_1^{**} K'(\lambda_0^{**}) - K(\lambda_1^{**}) = (1 - \lambda_1^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] > 0$$
   $$\Rightarrow K'(\lambda_0^{**}) > \frac{K(\lambda_1^{**})}{\lambda_1^{**}} = K'(\frac{\tilde{c}}{v_g})$$
   $$\Rightarrow \lambda_0^{**} > \frac{\tilde{c}}{v_g}$$
   $$\Rightarrow \tilde{c} < v_g\lambda_0^{**} < v_g\lambda_1^{**}$$

   $\square$

   We then compare the optimal sender surplus between the solution to (P_{2S_+}) and the solution to (P_{2S_0}), and show that the optimal S_0 strategy is always preferred to the optimal S_+ strategy when both types of strategy are feasible.

   Proposition 13. *Suppose $c \leq \tilde{c}$ and $\mu_0 < \widehat{\mu}_0$. The sender uses S_0 strategy when she provides information in both periods.*

---

[12] But the optimal 2-period strategy may be either S_+ or S_0 strategy.

Proof. $\forall \mu_0 < \widehat{\mu}_0$ such that S_0 (S_+) strategy is feasible, denote the optimal sender surplus by $\Pi_{S_0}(\mu_0)$ $(\Pi_{S_+}(\mu_0))$.

i) $\mu_0' := \frac{(v_g-c)(c-v_b\lambda_1^{**})+c}{v_g} \leq \mu_0 < \widehat{\mu}_0$

$\frac{(v_g-c)(c-v_b\lambda_1^{**})+c}{v_g} \leq \mu_0 \Leftrightarrow \widehat{\mu}_1(\mu_0) \leq c-v_b\lambda_1^{**}$. According to Proposition 11, $(\lambda_{0,S_+},\mu_{1,S_+}) = (\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, c - v_b\lambda_1^{**})$ gives $\Pi_{S_+}(\mu_0)$. $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c} \wedge c - v_b\lambda_1^{**}\right]$ in (P_{2S_0}'') and $\frac{(v_g-c)(c-v_b\lambda_1^{**})+c}{v_g} \leq \mu_0 \Leftrightarrow \frac{v_g\mu_0-c}{v_g-c} \geq c - v_b\lambda_1^{**}$. Consider $(\lambda_0,\mu_1) = (\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}, c - v_b\lambda_1^{**})$, which satisfies the costraints in (P_{2S_0}'') and is identical to $(\lambda_{0,S_+},\mu_{1,S_+})$. So, $\Pi_{S_0}(\mu_0) \geq \Pi_{S_+}(\mu_0)$.

ii) $\mu_0 < \mu_0'$

$\mu_0 < \mu_0' \Leftrightarrow \widehat{\mu}_1(\mu_0) > c - v_b\lambda_1^{**}$. According to Proposition 11, $(\lambda_{0,S_+},\mu_{1,S_+}) = (\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, \widehat{\mu}_1(\mu_0))$ gives $\Pi_{S_+}(\mu_0)$. One can verify that $\frac{v_g\mu_0'-c}{v_g-c} = c-v_b\lambda_1^{**}$, $\frac{\mu_0'-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c} = \frac{c}{v_g} \leq \lambda_0^{**}$. So, $\mu_0' \leq \widehat{\mu}_0$, which implies that $\lambda_{0,S_+}(\mu_0') \leq \lambda_0^{**}$.

First, consider μ_0 such that both S_0 and S_+ strategies are feasible. Let $\mu_1^u(\mu_0) := \frac{v_g\mu_0-c}{v_g-c}$.

$$\Pi_{S_+}(\mu_0) = - K(\lambda_{0,S_+}(\mu_0)) + p\lambda_{0,S_+}(\mu_0) + (1 - \lambda_{0,S_+}(\mu_0))[-K(\lambda_1^{**}) + p\lambda_1^{**}]$$

$$\Pi_{S_0}(\mu_0) \geq \underline{\Pi}_{S_0}(\mu_0) := - K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})\left[-K(\frac{\mu_1^u(\mu_0) - c}{-v_b}) + \frac{(\mu_1^u(\mu_0) - c)p}{-v_b}\right]$$

Lemma 10. $\frac{\mu_1^u(\mu_0)-c}{-v_b} \geq \lambda_{0,S_+}(\mu_0)$, $\forall \mu_0 < \mu_0'$.

Proof. $\frac{d}{d\mu_0}[\frac{\mu_1^u(\mu_0)-c}{-v_b}] = \frac{v_g}{-v_b(v_g-c)}$, $\frac{d}{d\mu_0}[\lambda_{0,S_+}(\mu_0)] = \frac{1}{-v_b(1-\lambda_1^{**})-c}$.

$$\frac{d}{d\mu_0}[\frac{\mu_1^u(\mu_0)-c}{-v_b}] \leq \frac{d}{d\mu_0}[\lambda_{0,S_+}(\mu_0)] \Leftrightarrow c(-2v_b - 1) \leq -v_bv_g\lambda_1^{**} \tag{$*$}$$

If $v_b \geq -1/2$, $(*)$ always holds. If $(-1 <) v_b < -1/2$, we have that $\frac{-v_b}{-2v_b-1} \geq 1 \Rightarrow c \leq v_g\lambda_1^{**} \leq \frac{-v_bv_g}{-2v_b-1}\lambda_1^{**} \Rightarrow (*)$ also holds. So, $\frac{d}{d\mu_0}[\frac{\mu_1^u(\mu_0)-c}{-v_b}] \leq \frac{d}{d\mu_0}[\lambda_{0,S_+}(\mu_0)]$, $\forall \mu_0 < \mu_0'$.

Note that $\frac{\mu_1^u(\mu_0')-c}{-v_b} = \lambda_1^{**} > \lambda_0^{**} \geq \lambda_{0,S_+}(\mu_0')$. This concludes the proof. $\square$

Now we calculate the increasing rate of the sender surplus as a function of μ_0:

$$\frac{d\Pi_{S_+}(\mu_0)}{d\mu_0} = \frac{K'(\lambda_{0,S_+}(\mu_0)) - K(\lambda_1^{**}) + \frac{cp}{v_b}}{v_b(1 - \lambda_1^{**}) + c} - \frac{p}{v_b}$$

$$\frac{d\underline{\Pi}_{S_0}(\mu_0)}{d\mu_0} := \frac{K'(\frac{\mu_1^u(\mu_0) - c}{-v_b})}{v_b} - \frac{p}{v_b}$$

$$\frac{d\Pi_{S_+}(\mu_0)}{d\mu_0} \geq \frac{d\underline{\Pi}_{S_0}(\mu_0)}{d\mu_0} \Leftrightarrow \frac{K'(\frac{\mu_1^u(\mu_0) - c}{-v_b})}{-v_b} + \frac{K(\lambda_1^{**}) - \frac{cp}{v_b}}{-v_b(1 - \lambda_1^{**}) - c} \geq \frac{K'(\lambda_{0,S_+}(\mu_0))}{-v_b(1 - \lambda_1^{**}) - c}$$

$$(\star)$$

$$c \leq \tilde{c} = v_g K'^{-1}\left[\frac{K(\lambda_1^{**})}{\lambda_1^{**}}\right]$$

$$\Leftrightarrow K(\lambda_1^{**}) \geq \lambda_1^{**} K'(\frac{c}{v_g})$$

$$\Rightarrow -v_b\left[K(\lambda_1^{**}) - \frac{cp}{v_b}\right] \geq (c - v_b\lambda_1^{**})K'(\frac{c}{v_g}) \left(c \leq v_g\lambda_1^{**} \Rightarrow K'(\frac{c}{v_g}) \leq K'(\lambda_1^{**}) = p\right)$$

$$\Rightarrow -v_b\left[K(\lambda_1^{**}) - \frac{cp}{v_b}\right] \geq (c - v_b\lambda_1^{**})K'(\lambda_{0,S_+}(\mu_0)) \left(\lambda_{0,S_+}(\mu_0) < \lambda_{0,S_+}(\mu_0') = \frac{c}{v_g}\right)$$

$$\Leftrightarrow \frac{K'(\lambda_{0,S_+}(\mu_0))}{-v_b} + \frac{K(\lambda_1^{**}) - \frac{cp}{v_b}}{-v_b(1 - \lambda_1^{**}) - c} \geq \frac{K'(\lambda_{0,S_+}(\mu_0))}{-v_b(1 - \lambda_1^{**}) - c}$$

$$\overset{Lemma10}{\Rightarrow} \frac{K'(\frac{\mu_1^u(\mu_0) - c}{-v_b})}{-v_b} + \frac{K(\lambda_1^{**}) - \frac{cp}{v_b}}{-v_b(1 - \lambda_1^{**}) - c} \geq \frac{K'(\lambda_{0,S_+}(\mu_0))}{-v_b(1 - \lambda_1^{**}) - c}$$

$$\overset{(\star)}{\Leftrightarrow} \frac{d\Pi_{S_+}(\mu_0)}{d\mu_0} \geq \frac{d\underline{\Pi}_{S_0}(\mu_0)}{d\mu_0}$$

One can verify that $\lambda_{0,S_+}(\mu_0') = \frac{c}{v_g}$, $\frac{\mu_1^u(\mu_0') - c}{-v_b} = \lambda_1^{**}$. So, $\Pi_{S_+}(\mu_0') = \underline{\Pi}_{S_0}(\mu_0')$. Therefore, $\underline{\Pi}_{S_0} \geq \Pi_{S_+}(\mu_0)$. $\Pi_{S_0}(\mu_0) \geq \underline{\Pi}_{S_0} \Rightarrow \Pi_{S_0}(\mu_0) \geq \Pi_{S_+}(\mu_0)$.

Now we show that the S_0 strategy is feasible whenever the S_+ strategy is feasible, which concludes the proof of Proposition 13.

Lemma 11. *Suppose $c < v_g\lambda_1^{**}$. For any μ_0 such that S_+ strategy is feasible, S_0 strategy is also feasible.*

Proof. Suppose there exists μ_0 such that S_+ strategy is feasible while S_0 strategy is not feasible. Then, $2c - v_b\lambda_1^{**} < \frac{2v_g - c}{(v_g)^2}$ and $\frac{(2 - \lambda_1^{**})c}{v_g(1 - \lambda_1^{**}) + c} < \frac{2v_g - c}{(v_g)^2}$, which is equivalent to $\lambda_1^{**} < \frac{c^2}{v_b(v_g)^2} + \frac{2c}{v_g}$ and $c > v_g\lambda_1^{**}$, which is not possible as we assumed that $c < v_g\lambda_1^{**}$. $\square$

$\square$

Proposition 13 tells us that we could limit our attention to S_0 strategy when $c \leq \tilde{c}$. Proposition 12 has characterized the optimal S_0 strategy. One can see from the characterization of the optimal strategy that the sender's payoff strictly increases in the prior.

$\square$

Proposition 14. (Comparative Statics W.r.t. the Prior When the Search Cost Is Intermediate) *When the search cost is intermediate, $v_g \lambda_0^{**} < c < v_g \lambda_1^{**}$, and the sender provides information in both periods. There exists $\mu_{2,+} \in [\frac{2v_g - c}{(v_g)^2}c, \widehat{\mu_0})$ and $\mu_{2,0} \in [\frac{2v_g - c}{(v_g)^2}c, \mu_{2,+}]$. The probability of a positive signal in the first period, λ_0^*, remains the same when $\mu_0 < \mu_{2,0}$ and strictly increases in the prior when $\mu_0 > \mu_{2,+}$. The probability of a positive signal in the second period, λ_1^*, strictly increases in the prior when $\mu_0 < \mu_{2,0}$ and remains the same when $\mu_0 > \mu_{2,+}$. The belief after observing a positive signal in the second period, $\bar{\mu}_1^*$, strictly decreases in the prior when $\mu_0 < \mu_{2,0}$ or $\mu_0 > \mu_{2,+}$. A positive signal always fully reveals the state in the first period, $\bar{\mu}_0^* \equiv 1$.*

When the expected receiver surplus in the second period is zero (S_0 strategy), the minimum amount of information for the receiver to search in the first period is already too high. Under the S_+ strategy, the receiver anticipates that the sender will provide favorable information in the second period, which relaxes the first-period participation constraint. Therefore, the receiver is willing to search even if the sender provides less information in the first period. This benefits the sender. However, the S_+ strategy has the disadvantage of inducing higher expected receiver search costs.

When the prior is low, the disadvantage of the S_+ strategy dominates the advantage. The sender prefers the S_0 strategy to the S_+ strategy. She faces information over-provision in the first period and under-provision in the second period. In the first period, she provides the minimum amount of information for the receiver to search. So, the probability of a positive signal in the first period, λ_0^*, does not depend on the prior. In the second period, she provides the maximum amount of information. More frequent positive signals are feasible when the prior is higher. Even if the receiver becomes less certain about the state being good after observing a positive signal, he would still search as long as the likelihood of receiving a positive signal and earning a strictly positive surplus increases. In equilibrium, the sender trades off the precision of a positive signal for frequency as the prior increases.

When the prior is high, the advantage of the S_+ strategy dominates the disadvantage. The sender prefers the S_+ strategy to the S_0 strategy. She faces information under-provision in the first period and no distortion in the second period. In the first period, she provides the maximum amount of feasible information, which strictly increases in the prior. In the

second period, she provides the optimal amount of information, which does not depend on the prior. When the prior increases, the participation constraint in the first period is relaxed. To persuade the receiver to search in the first period, the sender could provide information less favorable to the receiver in the second period. She trades off the precision of a positive signal in the second period for the frequency of positive signals in the first period. So, the belief after observing a positive signal, $\bar{\mu}_1^*$, strictly decreases in the prior. The optimal strategy is discontinuous when the sender switches the types of strategy (as illustrated in Figure 6).

Proof of Proposition 14. We compare the sender surplus between the solution to (P_{2S_+}) and the solution to (P_{2S_0}). The following result shows that the optimal S_+ strategy generates a strictly higher sender surplus than the optimal S_0 strategy when the prior is high.

Lemma 12. *If $v_g \lambda_0^{**} < c < v_g \lambda_1^{**}$, the receiver gets strictly positive surplus in the second period when μ_0 is close to $\widehat{\mu}_0$.*

Proof. According to Proposition 11, the sender achieves the benchmark payoff $-K(\lambda_0^{**}) + p\lambda_0^{**} + (1 - \lambda_0^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$ when $\mu_0 \to \widehat{\mu}_0$ by S_+ strategy. According to Lemma 8, the sender surplus of S_0 strategy is $\leq -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] < -K(\lambda_0^{**}) + p\lambda_0^{**} + (1 - \lambda_0^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$ as $\lambda_0^{**} < \frac{c}{v_g} = \lambda_0$. The difference of the benchmark sender surplus and the sender surplus of S_0 strategy is larger than a strictly positve constant. So, when $\mu_0 \to \widehat{\mu}_0$, S_+ strategy gives the sender strictly higher payoff. $v_g \lambda_0^{**} < c \Leftrightarrow \widehat{\mu}_1(\widehat{\mu}_0) > c - v_b \lambda_1^{**} \Rightarrow$ the receiver gets strictly positive surplus from S_+ strategy when $\mu_0 \to \widehat{\mu}_0$. Thus, the receiver gets strictly positive surplus in the second period in equilibrium when $\mu_0 \to \widehat{\mu}_0$. Continuity of the optimal strategy and sender surplus then implies that there exists a neighborhood $[\widehat{\mu}_0 - \delta, \widehat{\mu}_0]$ for some $\delta > 0$ such that the receiver gets strictly positive surplus in the second period in equilibrium when $\mu_0 \in [\widehat{\mu}_0 - \delta, \widehat{\mu}_0]$. $\square$

When $\frac{c}{v_g} \leq \mu_0 < \mu_{1,2}$, the sender provides information in one period. When $\mu_{1,2} \leq \mu_0 < \widehat{\mu}_0$, the sender provides information in both periods. Proposition 11, and Lemma 8 imply the explicit form of the optimal strategy. When $\mu_{1,2} \leq \mu_0 < \mu_{2,+}$, $(\lambda_0^*, \bar{\mu}_0^*) = (\frac{c}{v_g}, 1)$, $(\lambda_1^*, \bar{\mu}_1^*) = (\frac{v_g(\mu_0-c)-c(1-c)}{p(v_g-c)}, \frac{[v_g\mu_0-c]p}{v_g(\mu_0-c)-c(1-c)})$, $\mu_1^* = \frac{v_g\mu_0-c}{v_g-c} < c - v_b\lambda_1^{**}$. The receiver gets zero surplus at each period. When $\mu_{2,+} \leq \mu_0 < \widehat{\mu}_0$, $(\lambda_0^*, \bar{\mu}_0^*) = (\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, 1)$, $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}})$, $\mu_1^* = \widehat{\mu}_1(\mu_0) > c - v_b\lambda_1^{**}$, where $\widehat{\mu}_1(\mu_0) = \frac{2c-v_b\lambda_1^{**}-(1+c-v_b\lambda_1^{**}+v_b)\mu_0}{c-v_b-\mu_0}$. The receiver gets strictly positive surplus in the second period and zero total surplus. $\square$

Proof of Propostion 5. When the search cost is high, $v_g\lambda_1^{**} \leq c < \widehat{c}$, the optimal strategy of the sender is $(\lambda_t^*, \bar{\mu}_t^*) = (\frac{c}{v_g}, 1), t = 0, 1$ according to Proposition 4, which does not depend on η.

When the search cost is low, $c < \tilde{c} < v_g \lambda_0^{**}$. The boundary solution does not depend on η. Consider the interior solution to (P_{2S_0}), we have: $\eta \tilde{K}(\frac{\mu_1 - c}{-v_b}) + \frac{p - \mu_1}{p} \eta \tilde{K}'(\frac{\mu_1 - c}{-v_b}) - \eta \tilde{K}'(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}) + c = 0$. The LHS strictly increases in μ_1 and strictly decreases in η. So, $\eta \uparrow \Rightarrow \mu_1^* \uparrow \Rightarrow \lambda_1^* = \frac{\mu_1^* - c}{p} \uparrow, \lambda_0^* = \frac{\mu_0 - \mu_1^* - c}{p - \mu_1^*} \downarrow$. $\square$

Proposition 15. (Efficient Strategy for the Subgame) *At the second period, the social planner does not provide information when $\mu_1 < \mu_{0,1}$. When $\mu_1 \geq \mu_{0,1}$, the efficient signal fully reveals the state when a positive signal arrives, $\bar{\mu}_{1,e} = 1$; the probability of a positive signal, $\lambda_{1,e}$, depends on the search cost:*

1. *if $c \geq v_g \lambda_1^{**}$, then there exists a unique $\widehat{\widehat{c}} \in (v_g \lambda_1^{**}, \mu_1 v_g]$ such that the social planner does not provide information if $c > \widehat{\widehat{c}}$ and $\lambda_{1,e} = \frac{c}{v_g} \vee (\tilde{\lambda}_1 \wedge \mu_1)$ if $c \leq \widehat{\widehat{c}}$.*

2. *if $c \in [\mu_1 + v_b \lambda_1^{**}, v_g \lambda_1^{**})$, then $\lambda_{1,e} = \mu_1$.*

3. *if $c < \mu_1 + v_b \lambda_1^{**} \wedge v_g \lambda_1^{**}$, then $\lambda_{1,e} = \tilde{\lambda}_1 \wedge \mu_1$.*

Proof of Proposition 15. We first introduce a benchmark problem, in which the receiver is forced to participate and the social planner could generate any signal that fully reveals the state when a positive signal arrives. The social planner chooses the information structure to maximize total welfare. We will use the solution throughout the remaining section.

$$\max_{\lambda_1} -K(\lambda_1) + \lambda_1 \qquad (E_b)$$

Lemma 13. *The optimal solution to (E_b) exists and is unique. Denote it by $\tilde{\lambda}_1$. The objective function under $\tilde{\lambda}_1$ is strictly positive. $\tilde{\lambda}_1$ does not depend on the search cost c and $\tilde{\lambda}_1 > \lambda_1^{**}$, the solution to the payoff-maximizing benchmark problem.*

Proof. All the results follow from the same argument as the proof of Lemma 1 except $\tilde{\lambda}_1 > \lambda_1^{**}$. The F.O.C.'s imply $K'(\tilde{\lambda}_1) = 1 > p = K'(\lambda_1^{**}) \Rightarrow \tilde{\lambda}_1 > \lambda_1^{**}$. $\square$

As the social planner wants to maximize the total welfare, she always wants to make the precision of the signal, $\bar{\mu}_1$, as high as possible (subject to the feasibility constraint) given μ_1 and λ_1, to increase receiver surplus while holding sender surplus fixed. So, if the social planner provides information, then $\bar{\mu}_1 = \frac{\mu_1}{\lambda_1} \wedge 1$.

1. if $c \geq v_g \lambda_1^{**}$ (*i.e.* $\lambda_1^{**} \leq \frac{c}{v_g}$), then $TS = \begin{cases} -K(\lambda_1) + \mu_1 - c, & if \ \lambda_1 \geq \mu_1 \\ -K(\lambda_1) + \lambda_1 - c, & if \ \lambda_1 < \mu_1 \end{cases} \Rightarrow \lambda_{1,e} = \frac{c}{v_g} \vee (\tilde{\lambda}_1 \wedge \mu_1)$ when the social planner provides information. The social planner wiould

provide information if the total surplus is non-negative. Similar to the proof of Proposition 2, one can show that there exists a unique $\widehat{\widehat{c}} \in (v_g \lambda_1^{**}, \mu_1 v_g]$ such that the social planner does not provide information iff $c > \widehat{\widehat{c}}$. Moreover, if $\widehat{\widehat{c}} \geq (1-p)\tilde{\lambda}_1$, $\lambda_{1,e} = \frac{c}{1-p} \Rightarrow \bar{\mu}_{1,e} = 1 \Rightarrow TS = -K(\lambda_{1,e}) + p\lambda_{1,e} \Rightarrow \widehat{\widehat{c}} = \widehat{c}$.

2. if $c \in [\mu_1 - p\lambda_1^{**}, v_g \lambda_1^{**})$ $(i.e.\ \lambda_1^{**} \geq \frac{\mu_1 - c}{-v_b} > \frac{c}{1-p})$, then by the previous argument, $\lambda_{1,e} = \frac{c}{1-p} \vee (\tilde{\lambda}_1 \wedge \mu_1) = \mu_1$.

3. if $c < \mu_1 - p\lambda_1^{**} \wedge (1-p)\lambda_1^{**}$ $(i.e.\ \lambda_1^{**} \in \left(\frac{c}{1-p}, \frac{\mu_1 - c}{-v_b}\right))$, then the total welfare $TS =$
$$\begin{cases} -K(\lambda_1) + \mu_1 - c,\ if\ \lambda_1 \geq \mu_1 \\ -K(\lambda_1) + \lambda_1 - c,\ if\ \lambda_1 < \mu_1 \end{cases} \Rightarrow \lambda_{1,e} = \tilde{\lambda}_1 \wedge \mu_1.$$

$\square$

Proof of Proposition 6. When $(1-p)\tilde{\lambda}_1 \leq c \leq \widehat{\widehat{c}}$, the constrained program of the social planner is:

$$\max -K(\lambda_0) + \lambda_0 \bar{\mu}_0 - c + (1 - \lambda_0)\left[-K(\frac{c}{1-p}) + \frac{cp}{1-p}\right] \qquad (E_{2H})$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \geq \frac{c}{1-p}$$

Using similar methods of finding the optimal sender strategy in the main text, one can show that the solution to (E_{2H}) is $(\lambda_{0,e}, \lambda_{1,e}) = (\frac{c}{1-p}, \frac{c}{1-p})$. Therefore, there is no information distortion.

When $v_g \lambda_1^{**} \leq c < (1-p)\tilde{\lambda}_1 \wedge \widehat{\widehat{c}}$, the constrained program of the social planner is:

$$\max -K(\lambda_0) + \lambda_0 \bar{\mu}_0 - c + (1 - \lambda_0)\left[-K(\tilde{\lambda}_1 \wedge \mu_1) + \tilde{\lambda}_1 \wedge \mu_1 - c\right] \qquad (E_{2I})$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \geq \frac{c}{1-p}$$

Using similar methods of finding the optimal sender strategy in the main text, one can show that the solution to (E_{2I}) is $(\lambda_{0,e}, \lambda_{1,e}) = (\frac{c}{1-p}, \frac{c}{1-p})$. Therefore, there is no information distortion.

When $c < v_g \lambda_1^{**}$, one can see that we could restrict μ_1 to be less than or equal to $\tilde{\lambda}_1$ without loss of generality. The constrained program of the social planner is:

$$\max -K(\lambda_0) + \lambda_0 \bar{\mu}_0 - c + (1 - \lambda_0)\left[-K(\mu_1) + \mu_1 - c\right] \qquad (E_{2S_0})$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \geq \frac{c}{1-p}$$

Using similar methods of finding the optimal sender strategy in the main text, one can show that the solution to (E_{2S_0}) is $(\lambda_{0,e}, \lambda_{1,e}) = (\frac{\mu_0 - \frac{c}{1-p} - c}{p - \frac{c}{1-p}}, \frac{c}{1-p})$. Therefore, the sender provides too much information (relative to the efficient solution) in the second period and too little information in the first period. $\qquad\square$

Proof of Proposition 7. One can see that if the expected surplus of the receiver in the second period is 0 in the optimal solution to (P_{dc}), then $(\lambda_1^*, \bar\mu_1^*)$ solves (P_1). Otherwise, the sender can strictly increase the payoff by using the same $(\lambda_0^*, \bar\mu_0^*)$ and replacing $(\lambda_1^*, \bar\mu_1^*)$ by the optimal solution to (P_1), holding the same μ_1. Hence, if dynamic commitment power strictly increases the sender surplus, the solution to (P_{dc}) must satisfy: $\mathbb{E}[\text{receiver surplus at } t = 1] = (1 - \lambda_0)[\lambda_1(\bar\mu_1 + v_b) - c] > 0$. Denote the optimal sender surplus when the sender does not have dynamic commitment power by Π_{wo}. The corresponding optimal strategy of the sender is $(\lambda_t^*, \bar\mu_t^*) = (\frac{c}{v_g}, 1)$ according to Proposition 4. Consider the following strategy: $(\lambda_0, \bar\mu_0, \lambda_1, \bar\mu_1) = (\frac{c - \delta v_g}{(1-\delta)v_g}, 1, \frac{c}{v_g} + \delta, 1)$. Denote the corresponding sender surplus by $\Pi(\delta)$. One can verify that it is feasible when $\delta > 0$ is small and the sender has dynamic commitment power, and leads to a payoff no larger than the optimal sender surplus with dynamic commitment (denote it by Π_w).

$$\Pi(\delta) = - K\left[\frac{c - \delta v_g}{(1-\delta)v_g}\right] + p\frac{c - \delta v_g}{(1-\delta)v_g} + \left[1 - \frac{c - \delta v_g}{(1-\delta)v_g}\right]\left[-K(\frac{c}{v_g} + \delta) + \frac{cp}{v_g} + \delta p\right]$$

$$\frac{d\Pi(\delta)}{d\delta} = \frac{v_g - c}{(1-\delta)^2 v_g}\left[I_1(\delta) + I_2(\delta)\right]$$

$$, \text{ where } I_1(\delta) = K'\left[\frac{c - \delta v_g}{(1-\delta)v_g}\right] - (1-\delta)K'(\frac{cp}{v_g} + \delta)$$

$$I_2(\delta) = \frac{cp}{v_g} - K(\frac{c}{v_g} + \delta)$$

$I_1(\delta) \to 0$, $I_2(\delta) \to \frac{cp}{v_g} - K(\frac{c}{v_g}) > 0$ as $\delta \to 0$. Therefore, $\exists \widehat\delta > 0$ s.t. $\frac{d\Pi(\delta)}{d\delta} > 0$, $\forall \delta \in (0, \widehat\delta]$. Since $\Pi(\delta)$ is continuous in δ and $\Pi(0) = \Pi_{wo}$, we have $\Pi_w \geq \Pi(\delta) > \Pi_{wo}$, $\forall \delta \in (0, \widehat\delta]$.

We now show that the benefit of dynamic commitment power vanishes as the search cost approaches zero in several steps. Denote the optimal sender surplus when the sender has (does not have) dynamic commitment power and the search cost is c by $\Pi_{w,c}$ ($\Pi_{wo,c}$). Note that $\Pi_{w,c} \geq \Pi_{wo,c}, \forall c \geq 0$.

Lemma 14. *When the search cost is zero, the optimal sender surpluses with and without dynamic commitment power are the same, $\Pi_{w,0} = \Pi_{wo,0}$.*

Proof. Suppose the sender provides information in both periods. When the search cost is 0

52

and the sender uses one-shot experiments, one can see that (IR_0) and (IR_1) are equivalent to $\bar{\mu}_0 \geq -v_b$ and $\bar{\mu}_1 \geq -v_b$. Therefore, even if the receiver obtains strictly positive surplus in the second period, the first-period participation constraint is not relaxed. Hence, the second-period strategy of the sender maximizes her payoff of the subgame in the solution to (P_{dc}), and thus satisfies the constraints of (P_2). To show that dynamic commitment power does not improve the sender's payoff when $c = 0$, we just need to show that iterative signals are not optimal when the sender has dynamic commitment power.[13] When $c = 0 < v_g \lambda_1^{**}$ and $\mu_1 > c - v_b \lambda_1^{**} = -v_b \lambda_1^{**}$, the optimal strategy with and without dynamic commitment power coincides, as $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}} \wedge 1)$. In the proof of Proposition 1, we have shown that iterative signals are not optimal. So, we just need to show that iterative signals are not optimal when $\mu_1 \leq -v_b \lambda_1^{**}$.

If $\mu_0 \geq -v_b \lambda_1^{**}$, then $\mu_1 = \bar{\mu}_0 > \mu_0 \geq -v_b \lambda_1^{**}$. So, iterative signals are not optimal.

If $\mu_0 < -v_b \lambda_1^{**}$, we have:

$$\Pi_1(\mu_0) = -K\left(\frac{\mu_0}{-v_b}\right) + \frac{\mu_0 p}{-v_b}$$

$$\Pi_{iter}(\mu_0) = \max_{\lambda_0, \mu_1} -K(\lambda_0) + \lambda_0 \left[-K\left(\frac{\mu_1}{-v_b}\right) + \frac{\mu_1 p}{-v_b}\right]$$

$$\text{s.t. } (IR_{0,iter}), (IR_{1,iter}), (F_0), (F_1), \mu_1 = \bar{\mu}_0$$

Denote the optimal solution when the sender uses iterative signals by $(\tilde{\lambda}_0, \tilde{\mu}_1)$. $\Pi_{iter}(\mu_0) \leq \Pi_1(\mu_0) \Leftrightarrow -K(\tilde{\lambda}_0) + \tilde{\lambda}_0 \left[-K\left(\frac{\tilde{\mu}_1}{-v_b}\right) + \frac{\tilde{\mu}_1 p}{-v_b}\right] \leq -K\left(\frac{\mu_0}{-v_b}\right) + \frac{\mu_0 p}{-v_b}$. To show that iterative signals are not optimal, it suffices to show that $-\tilde{\lambda}_0 K\left(\frac{\tilde{\mu}_1}{-v_b}\right) + \frac{\tilde{\lambda}_0 \tilde{\mu}_1 p}{-v_b} \leq -K\left(\frac{\mu_0}{-v_b}\right) + \frac{\mu_0 p}{-v_b}$. Strict convexity of $K(\cdot) \Rightarrow \tilde{\lambda}_0 K\left(\frac{\tilde{\mu}_1}{-v_b}\right) = \tilde{\lambda}_0 K\left(\frac{\tilde{\mu}_1}{-v_b}\right) + (1 - \tilde{\lambda}_0) K(0) \geq K\left(\frac{\tilde{\lambda}_0 \tilde{\mu}_1}{-v_b}\right)$. Thus, $-\tilde{\lambda}_0 K\left(\frac{\tilde{\mu}_1}{-v_b}\right) \leq -K\left(\frac{\tilde{\lambda}_0 \tilde{\mu}_1}{-v_b}\right)$. It suffices to show that $-K\left(\frac{\tilde{\lambda}_0 \tilde{\mu}_1}{-v_b}\right) + \frac{\tilde{\lambda}_0 \tilde{\mu}_1 p}{-v_b} \leq -K\left(\frac{\mu_0}{-v_b}\right) + \frac{\mu_0 p}{-v_b}$, which hold because $(F_0) \Rightarrow \tilde{\lambda}_0 \tilde{\mu}_1 \leq \mu_0$ and we know from the F.O.C. that $-K(\lambda) + p\lambda$ strictly increases in λ when $\lambda < \lambda_1^{**}$ (here, $\lambda \leq \frac{\mu_0}{-v_b} < \lambda_1^{**}$). $\qquad \square$

We now make the following observation. Since $\Pi_{w,c} \geq \Pi_{wo,c}, \forall c \geq 0$ and $\Pi_{w,0} = \Pi_{wo,0}$, we must have $\Pi_{wo,c} \to \Pi_{w,c}$ as $c \to 0$ if $\Pi_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$.[14] The next result confirms that it is indeed the case and thus finishes the proof.

Lemma 15. $\Pi_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$.

Proof. Proposition 13 shows that the optimal strategy is S_0 strategy when the search cost

---

[13] We have shown in Proposition 1 that iterative signals are not optimal when the sender does not have dynamic commitment power.

[14] One can easily show this observation formally by the triangle inequality.

is low. So, according to Lemma 6 and 7, for any c small enough, the sender's problem is:

$$\Pi_{wo,c} = \max_{\mu_1} -K(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}) + p\,\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1} + (1 - \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1})\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right]$$

$$(P'''_{2S_0})$$

$$\text{s.t. } \mu_1 \in \left[\frac{c}{v_g}, \frac{v_g \mu_0 - c}{v_g - c}\right]$$

Denote the solution when the search cost is 0 by $\mu^*_{1,0}$. Define $\mu_{1,c}$ to be the closest value to $\mu^*_{1,0}$ among $\left[\frac{c}{v_g}, \frac{v_g \mu_0 - c}{v_g - c}\right]$ and denote the corresponding sender surplus by $\underline{\Pi}_{wo,c}$. One can see that $\underline{\Pi}_{wo,c} \leq \Pi_{wo,c}$. Since $\mu^*_{1,0} \in \left[0, \frac{v_g \mu_0}{v_g}\right]$, we have $\mu_{1,c} \to \mu^*_{1,0}$ as $c \to 0$. Therefore, $\underline{\Pi}_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$. Since $\underline{\Pi}_{wo,c} \leq \Pi_{wo,c} \leq \Pi_{wo,0}$, we also have $\Pi_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$. $\qquad \square$

$\square$

Proof of Proposition 8. When $c \geq \hat{c}$, the sender does not provide information for any δ.

(1) $v_g \lambda^{**}_1 \leq c < \hat{c}$

By the same argument as the proof of Proposition 4, one can see that the optimal strategy of the sender does not depend on δ. For low prior, $\mu_0 \in [\frac{c}{v_g}, \frac{2v_g - c}{(v_g)^2}c)$, the sender provides information in one period, $(\lambda^*, \bar{\mu}^*) = (\frac{c}{v_g}, 1)$. For high prior, $\mu_0 \geq \frac{2v_g - c}{(v_g)^2}c$, the sender provides information in both periods, $(\lambda^*_t, \bar{\mu}_t^*) = (\frac{c}{v_g}, 1), t = 0, 1$.

(2) $c \leq \tilde{c}$

The sender's problem can be divided into 2 cases:

$$\max -K(\lambda_0) + p\lambda_0 + \delta(1 - \lambda_0)\left[-K(\lambda^{**}_1) + p\lambda^{**}_1\right] \qquad (P^\delta_{2S_+})$$

$$\text{s.t. } (IR_{0,\delta}), (F_0), \mu_1 \in [c - v_b \lambda^{**}_1, \lambda^{**}_1]$$

$$\max -K(\lambda_0) + p\lambda_0 + \delta(1 - \lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] \qquad (P^\delta_{2S_0})$$

$$\text{s.t. } (IR_{0,\delta}), (F_0), \mu_1 \in [\frac{c}{v_g}, c - v_b \lambda^{**}_1]$$

Consider the solution to $(P^\delta_{2S_0})$. By the same argument as the proof of Lemma 6, $(P^\delta_{2S_0})$

54

is equivalent to

$$\max -K(\lambda_0) + p\lambda_0 + \delta(1 - \lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] \qquad (P_{2S_0}^{\delta'})$$

$$\text{s.t. } \lambda_0 \in \left[\frac{\mu_0 - \mu_1}{1 - \mu_1}, \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}\right]$$

$$\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right]$$

By arguments similar to the proof of Lemma 7, $\lambda_0 \leq \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}$ is binding for $\mu_0 < \widehat{\mu}_0$. Suppose that μ_1's constraints are not binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + \delta(1 - \lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] + \eta\left(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1} - \lambda_0\right)$ s.t. $\eta \geq 0, \eta\left(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1} - \lambda_0\right) = 0$. F.O.C. $\Rightarrow$

$$\eta = \delta(p - \mu_1) - \delta\frac{p - \mu_1}{p}K'(\frac{\mu_1 - c}{-v_b})$$

$$= -K'(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}) + p + \delta\left[K(\frac{\mu_1 - c}{-v_b}) - \mu_1 + c\right]$$

$$\Rightarrow \delta\left[K(\frac{\mu_1 - c}{-v_b}) + \frac{p - \mu_1}{p}K'(\frac{\mu_1 - c}{-v_b}) - p + c\right] - K'(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}) + p = 0 \qquad (*_\delta)$$

The sum of the first two terms of the LHS of $(*_\delta)$ strictly increases in δ, and the LHS of $(*_\delta)$ strictly increases in μ_1. So, the optimal λ_0 strictly decreases in δ; the optimal μ_1 and λ_1 strictly increase in δ. When one of μ_1's constraints is binding, λ_0 and λ_1 does not depend on δ.

We finish the proof of this case by showing that S_0 strategy dominates S_+ strategy. Denote the sender surplus of the optimal S_0 (S_+) strategy when the discount factor is $\delta \in (0, 1)$ by $\Pi_{S_0}(\delta)$ ($\Pi_{S_+}(\delta)$) and denote the corresponding optimal S_0 (S_+) strategy at

time t by $\lambda_{t,S_0}(\delta)$ $(\lambda_{t,S_+}(\delta))$. We have:

$$
\begin{aligned}
\Pi_{S_0}(\delta) =& - K(\lambda_{0,S_0}(\delta)) + p\lambda_{0,S_0}(\delta) + \delta(1 - \lambda_{0,S_0}(\delta)) \left[-K(\lambda_{1,S_0}(\delta)) + p\lambda_{1,S_0}(\delta)\right] \\
\geq& - K(\lambda_{0,S_0}(1)) + p\lambda_{0,S_0}(1) + \delta(1 - \lambda_{0,S_0}(1)) \left[-K(\lambda_{1,S_0}(1)) + p\lambda_{1,S_0}(1)\right] \\
=& - K(\lambda_{0,S_0}(1)) + p\lambda_{0,S_0}(1) + (1 - \lambda_{0,S_0}(1)) \left[-K(\lambda_{1,S_0}(1)) + p\lambda_{1,S_0}(1)\right] \\
& - (1-\delta)(1 - \lambda_{0,S_0}(1)) \left[-K(\lambda_{1,S_0}(1)) + p\lambda_{1,S_0}(1)\right] \\
\overset{(\dagger)}{\geq}& - K(\lambda_{0,S_+}(1)) + p\lambda_{0,S_+}(1) + (1 - \lambda_{0,S_+}(1)) \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \\
& - (1-\delta)(1 - \lambda_{0,S_+}(1)) \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \\
=& - K(\lambda_{0,S_+}(1)) + p\lambda_{0,S_+}(1) + \delta(1 - \lambda_{0,S_+}(1)) \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \\
\geq& - K(\lambda_{0,S_+}(\delta)) + p\lambda_{0,S_+}(\delta) + \delta(1 - \lambda_{0,S_+}(\delta)) \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \\
=& \ \Pi_{S_+}(\delta)
\end{aligned}
$$

, where the inequality $(\dagger)$ holds because Proposition 13 implies that $-K(\lambda_{0,S_0}(1)) + p\lambda_{0,S_0}(1) + (1 - \lambda_{0,S_0}(1)) \left[-K(\lambda_{1,S_0}(1)) + p\lambda_{1,S_0}(1)\right] \geq -K(\lambda_{0,S_+}(1)) + p\lambda_{0,S_+}(1) + (1 - \lambda_{0,S_+}(1)) \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$ and we have $\lambda_{0,S_0}(1) \geq \lambda_{0,S_+}(1)$, $-K(\lambda_{1,S_0}(1)) + p\lambda_{1,S_0}(1) \leq -K(\lambda_1^{**}) + p\lambda_1^{**}$.

$\square$

Proof of Proposition 9.

(1) High search cost $v_g\lambda_1^{**} \leq c < \widehat{c}$

Because the receiver searches for a finite number of periods, we can do backward induction. Suppose the sender provides information for k periods.

In the last period, $t = k - 1$, $(\lambda_{k-1}^*, \bar{\mu}_{k-1}^*) = (c/v_g, 1)$ according to Proposition 2. The receiver gets zero expected surplus.

Suppose the optimal experiment in time t is one-shot and $(\lambda_t^*, \bar{\mu}_t^*) = (c/v_g, 1)$ for any $t \geq t'$ $(t' \geq 1)$, we want to show that the optimal experiment in time $t = t'-1$ is one-shot and $(\lambda_{t'-1}^*, \bar{\mu}_{t'-1}^*) = (c/v_g, 1)$.

Because the receiver's expected payoff during time $t \geq t'$ is zero, the optimal experiment in time $t' - 1$ is one-shot - the receiver takes action G right after observing a positive signal. Suppose not, then the receiver searches for more information regardless of the signal realization (clearly, he will not search and take action G regardless of the signal realization). But then, the receiver's expected surplus will be strictly negative because search is costly and he does not gain any utility in $t = t' - 1$. Hence, the optimal

experiment in time $t' - 1$ is one-shot. We just need to show that $(\lambda_{t'-1}^*, \bar{\mu}_{t'-1}^*) = (c/v_g, 1)$. The sender's problem in $t = t' - 1$ is:

$$\max_{\lambda_t, \bar{\mu}_t} - K(\lambda_t) + p\lambda_t + (1 - \lambda_t)V_t$$

$$\text{s.t. } \lambda_t(\bar{\mu}_t + v_b) \geq c \qquad (IR_t)$$

$$\lambda_t\bar{\mu}_t + (1 - \lambda_t)\underline{\mu}_t = \mu_t \qquad (F_t)$$

$$(\lambda_t, \bar{\mu}_t) = (c/v_g, 1), \ \forall t \geq t'$$

$$V_t = \sum_{i=0}^{k-t-1} (1 - \frac{c}{v_g})^i \left[\frac{cp}{v_g} - K(\frac{c}{v_g}) \right]$$

Denote the optimal λ_t without constraints by $\lambda_{t,H}^{**}$. $\lambda_{t,H}^{**} = \operatorname{argmax}_{\lambda_t} -K(\lambda_t) + p\lambda_t + (1 - \lambda_t)V_t$. The F.O.C. of $-K(\lambda_t) + p\lambda_t + (1 - \lambda_t)V_t \Rightarrow K'(\lambda_{t,H}^{**}) = p - V_t < p = K'(\lambda_1^{**}) \Rightarrow \lambda_{t,H}^{**} < \lambda_1^{**} \leq c/v_g$. So, $(\lambda_t^*, \bar{\mu}_t^*) = (c/v_g, 1)$.

Therefore, the induction shows that the sender provides one-shot experiment $(\lambda_t^*, \bar{\mu}_t^*) = (c/v_g, 1)$ in each period, $t = 0, 1, ..., k - 1$.

The sender's expected payoff of providing k periods of such information is: $\sum_{i=0}^{k-1}(1 - \frac{c}{v_g})^i \left[\frac{cp}{v_g} - K(\frac{c}{v_g}) \right]$, which increases in k. Thus, the sender will provide as many periods of information as possible. Now we characterize the maximum number of periods.

Denote the initial belief at the beginning of period t by $\mu_t := \underline{\mu}_{t-1}$. The feasibility costraint (F_t) and $(\lambda_t^*, \bar{\mu}_t^*) = (c/v_g, 1)$ imply that $\mu_t = \frac{\mu_{t-1} - \frac{c}{v_g}}{1 - \frac{c}{v_g}}$. By induction, one can show that $\mu_k = \frac{\mu_0 - 1 + (1 - \frac{c}{v_g})^k}{(1 - \frac{c}{v_g})^k}$. For it to be feasible to provide information in k periods, we need $\mu_{k-1} \geq c/v_g \Leftrightarrow k \leq \frac{ln(1-\mu_0)}{ln(1-c/v_g)}$. Hence, the maximum number of periods is $\lfloor \frac{ln(1-\mu_0)}{ln(1-c/v_g)} \rfloor$.

(2) Low search cost

The receiver needs to decide between G and B at the end of the game. We first derive an upper bound on the probability that the receiver decides on G under any feasible information structure.

Lemma 16. *The probability that the receiver takes action G eventually is bounded from above by* $-\frac{\mu_0}{v_b}$.

Proof. $\lambda_t\bar{\mu}_t + (1 - \lambda_t)\underline{\mu}_t = \mu_t, \forall t \Rightarrow \mu_0 \geq \sum_{t=0}^{+\infty} \lambda_t\bar{\mu}_t$

□

Corollary 1. *The sender's payoff when the persuasion cost is zero is bounded from above by* $-\frac{\mu_0 p}{v_b}$.

Now we show that the sender can achieve that payoff as the search cost vanishes, given any persuasion costs satisfying assumption 1. This means that she could obtain the equilibrium payoff as if the persuasion cost were zero.

Consider the following strategy: Given a search cost c, the sender provides the same one-shot experiments for T consecutive periods $(t = 0, 1, ..., T-1)$, where $(\lambda_t, \bar{\mu}_t) = (\lambda, \bar{\mu}) = (\sqrt{c}, -v_b + \sqrt{c})$ and $T = \frac{ln\left(1 - \frac{\mu_0}{-v_b + \sqrt{c}}\right)}{ln(1 - \sqrt{c})}$. In each period, the receiver's expected payoff from searching is $\lambda_t(\bar{\mu}_t + v_b) - c = 0$. So, the receiver will keep searching if he observes a negative signal, except in the last period. By setting $\underline{\mu}_{T-1} = 0$, one can verify that the mean-preserving property of the belief is satisfied, and that the variables are well-defined for c small.

The probability that the receiver takes action G eventually is $\sum_{t=0}^{T-1} \lambda(1-\lambda)^t = 1 - (1-\lambda)^T = \frac{\mu_0}{-v_b + \sqrt{c}} \to -\frac{\mu_0}{v_b}$ as $c \to 0$.

The sender only incurs the persuasion cost if the receiver has not received a good signal. The expected total persuasion cost of the sender is bounded from above by the costs of always providing the information in T periods, which is $TK(\sqrt{c})$.

$$\lim_{c \to 0} TK(\sqrt{c}) = \lim_{c \to 0} \frac{ln\left(1 - \frac{\mu_0}{-v_b + \sqrt{c}}\right)}{ln(1 - \sqrt{c})} K(\sqrt{c}) = ln\left(1 + \frac{\mu_0}{v_b}\right) \lim_{c \to 0} \frac{K(\sqrt{c})}{ln(1 - \sqrt{c})} \overset{\text{L'Hospital's rule}}{=} 0$$

Hence, the expected sender's payoff approaches $-\frac{\mu_0 p}{v_b}$ as $c \to 0$.

$\square$